

Lecture Notes
Combinatorics

Lecture by Maria Axenovich and Torsten Ueckerdt (KIT)
Problem Classes by Jonathan Rollin (KIT)

Lecture Notes by Stefan Walzer (TU Ilmenau)

Last updated: May 30, 2017

Contents

0	What is Combinatorics?	4
1	Permutations and Combinations	9
1.1	Basic Counting Principles	9
1.1.1	Addition Principle	10
1.1.2	Multiplication Principle	10
1.1.3	Subtraction Principle	11
1.1.4	Bijection Principle	11
1.1.5	Pigeonhole Principle	11
1.1.6	Double counting	11
1.2	Ordered Arrangements – Strings, Maps and Products	12
1.2.1	Permutations	13
1.3	Unordered Arrangements – Combinations, Subsets and Multisets	15
1.4	Multinomial Coefficients	16
1.5	The Twelvelfold Way – Balls in Boxes	22
1.5.1	$U \rightarrow L$: n Unlabeled Balls in k Labeled Boxes	22
1.5.2	$L \rightarrow U$: n Labeled Balls in k Unlabeled Boxes	25
1.5.3	$L \rightarrow L$: n Labeled Balls in k Labeled Boxes	27
1.5.4	$U \rightarrow U$: n Unlabeled Balls in k Unlabeled Boxes	28
1.5.5	Summary: The Twelvelfold Way	30
1.6	Binomial Coefficients – Examples and Identities	30
1.7	Permutations of Sets	34
1.7.1	Generating Permutations	36
1.7.2	Cycle Decompositions	37
1.7.3	Transpositions	39
1.7.4	Derangements	42
2	Inclusion-Exclusion-Principle and Möbius Inversion	45
2.1	The Inclusion-Exclusion Principle	45
2.1.1	Applications	47
2.1.2	Stronger Version of PIE	52
2.2	Möbius Inversion Formula	53
3	Generating Functions	58
3.1	Newton’s Binomial Theorem	62
3.2	Exponential Generating Functions	64
3.3	Recurrence Relations	67
3.3.1	Special Solution $a_n = x^n$ and the Characteristic Polynomial	70
3.3.2	Advancement Operator	76
3.3.3	Non-homogeneous Recurrences	80
3.3.4	Solving Recurrences using Generating Functions	82

4	Partitions	85
4.1	Partitioning $[n]$ – the set on n elements	85
4.1.1	Non-Crossing Partitions	86
4.2	Partitioning n – the natural number	87
4.3	Young Tableau	93
4.3.1	Counting Tableaux	99
4.3.2	Counting Tableaux of the Same Shape	100
5	Partially Ordered Sets	107
5.1	Subposets, Extensions and Dimension	111
5.2	Capturing Posets between two Lines	117
5.3	Sets of Sets and Multisets – Lattices	123
5.3.1	Symmetric Chain Partition	126
5.4	General Lattices	130
6	Designs	132
6.1	(Non-)Existence of Designs	133
6.2	Construction of Designs	135
6.3	Projective Planes	137
6.4	Steiner Triple Systems	138
6.5	Resolvable Designs	139
6.6	Latin Squares	140

What is Combinatorics?

Combinatorics is a young field of mathematics, starting to be an independent branch only in the 20th century. However, combinatorial methods and problems have been around ever since. Many combinatorial problems look entertaining or aesthetically pleasing and indeed one can say that roots of combinatorics lie in mathematical recreations and games. Nonetheless, this field has grown to be of great importance in today's world, not only because of its use for other fields like physical sciences, social sciences, biological sciences, information theory and computer science.

Combinatorics is concerned with:

- Arrangements of elements in a set into patterns satisfying specific rules, generally referred to as *discrete structures*. Here “discrete” (as opposed to continuous) typically also means finite, although we will consider some infinite structures as well.
- The existence, enumeration, analysis and optimization of discrete structures.
- Interconnections, generalizations- and specialization-relations between several discrete structures.

Existence: We want to arrange elements in a set into patterns satisfying certain rules. Is this possible? Under which conditions is it possible? What are necessary, what sufficient conditions? How do we find such an arrangement?

Enumeration: Assume certain arrangements are possible. *How many* such arrangements exist? Can we say “there are at least this many”, “at most this many” or “exactly this many”? How do we generate all arrangements efficiently?

Classification: Assume there are many arrangements. Do some of these arrangements differ from others in a particular way? Is there a natural partition of all arrangements into specific classes?

Meta-Structure: Do the arrangements even carry a natural underlying structure, e.g., some ordering? When are two arrangements closer to each other or more similar than some other pair of arrangements? Are different classes of arrangements in a particular relation?

Optimization: Assume some arrangements differ from others according to some measurement. Can we find or characterize the arrangements with maximum or minimum measure, i.e. the “best” or “worst” arrangements?

Interconnections: Assume a discrete structure has some properties (number of arrangements, ...) that match with another discrete structure. Can we specify a concrete connection between these structures? If this other structure is well-known, can we draw conclusions about our structure at hand?

We will give some life to this abstract list of tasks in the context of the following example.

Example (Dimer Problem). Consider a generalized chessboard of size $m \times n$ (m rows and n columns). We want to cover it perfectly with *dominoes* of size 2×1 or with generalized dominoes – called *polyominoes* – of size $k \times 1$. That means we want to put dominoes (or polyominoes) horizontally or vertically onto the board such that every square of the board is covered and no two dominoes (or polyominoes) overlap. A perfect covering is also called *tiling*. Consider Figure 1 for an example.

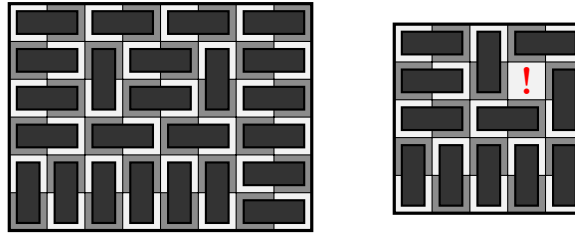


Figure 1: The 6×8 board can be tiled with 24 dominoes. The 5×5 board cannot be tiled with dominoes.

Existence

If you look at Figure 1, you may notice that whenever m and n are both odd (in the Figure they were both 5), then the board has an odd number of squares and a tiling with dominoes is not possible. If, on the other hand, m is even or n is even, a tiling can easily be found. We will generalize this observation for polyominoes:

Claim. An $m \times n$ board can be tiled with polyominoes of size $1 \times k$ if and only if k divides m or n .

Proof. “ \Leftarrow ” If k divides m , it is easy to construct a tiling: Just cover every column with m/k vertical polyominoes. Similarly, if k divides n , cover every row using n/k horizontal polyominoes.

“ \Rightarrow ” Assume k divides neither m nor n (but note that k could still divide the product $m \cdot n$). We need to show that no tiling is possible. We write $m = s_1k + r_1$, $n = s_2k + r_2$ for appropriate $s_1, s_2, r_1, r_2 \in \mathbb{N}$ and $0 < r_1, r_2 < k$. Without loss of generality, assume $r_1 \leq r_2$ (the argument is similar if $r_2 < r_1$). Consider the colouring of the $m \times n$ board with k colours as shown in Figure 2.

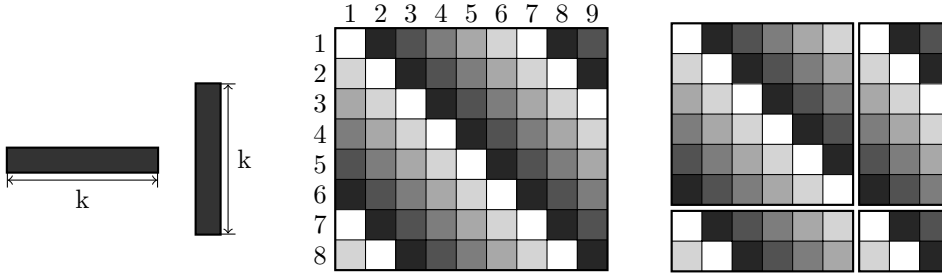


Figure 2: Our polyominoes have size $k \times 1$. We use k colours (1 = white, k = black) to colour the $m \times n$ board (here: $k = 6$, $m = 8$, $n = 9$). Cutting the board at coordinates that are multiples of k divides the board into several chunks. All chunks have the same number of squares of each color, except for the bottom right chunk where there are more squares of color 1 (here: white) than of color 2 (here: light gray).

Formally, the colour of the square (i, j) is defined to be $((i - j) \bmod k) + 1$. Any polyomino of size $k \times 1$ that is placed on the board will cover exactly one square of each colour. However, there are more squares of colour 1 than of colour 2, which shows that no tiling with $k \times 1$ dominoes is possible.

Indeed, for the number of squares coloured with 1 and 2 we have:

$$\begin{aligned} \# \text{ squares coloured with 1} &= ks_1s_2 + s_1r_2 + s_2r_1 + r_2 \\ \# \text{ squares coloured with 2} &= ks_1s_2 + s_1r_2 + s_2r_1 + r_2 - 1 \end{aligned} \quad \square$$

Now that the existence of tilings is answered for rectangular boards, we may be inclined to consider other types of boards as well:

Claim (Mutilated Chessboard). The $n \times n$ board with bottom-left and top-right square removed (see Figure 3) cannot be tiled with (regular) dominoes.

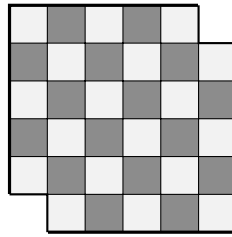


Figure 3: A “mutilated” 6×6 board. The missing corners have the same colour.

Proof. If n is odd, then the total number of squares is odd and clearly no tiling can exist. If n is even, consider the usual chessboard-colouring: In it, the missing squares are of the same colour, say black. Since there was an equal number of black and white squares in the non-mutilated board, there are now two more white squares than black squares. Since dominoes always cover exactly one black and one white square, no tiling can exist. \square

Other ways of pruning the board have been studied, but we will not consider them here.

Enumeration

A general formula to determine the number of ways an $m \times n$ board can be tiled with dominoes is known. For an $2m \times 2n$ board the following formula is due to Temperly and Fisher [TF61] and independently Kasteleyn [Kas61]

$$4^{mn} \prod_{i=1}^m \prod_{j=1}^n \left(\cos^2 \frac{i\pi}{2m+1} + \cos^2 \frac{j\pi}{2n+1} \right).$$

The special case of an 8×8 board is already non-trivial:

Theorem (Fischer 1961). *There are $2^4 \cdot 17^2 \cdot 53^2 = 12,988,816$ ways to tile the 8×8 board with dominoes.*

Classification

Consider tilings of the 4×4 board with dominoes. For some of these tilings there is a vertical line through the board that does not cut through any domino. Call such a line a *vertical cut*. In the same way we define *horizontal cuts*.

As it turns out, for every tiling of the 4×4 board at least one cut exists, possibly several (try this for yourself!).

Hence the set \mathcal{T} of all tilings can be partitioned into

$$\begin{aligned} \mathcal{T}_1 &= \{T \mid T \text{ allows a horizontal cut but no vertical cut}\}, \\ \mathcal{T}_2 &= \{T \mid T \text{ allows a vertical cut but no horizontal cut}\}, \\ \mathcal{T}_3 &= \{T \mid T \text{ allows both a horizontal and a vertical cut}\}. \end{aligned}$$

Figure 4 shows one tiling for each of these three classes.

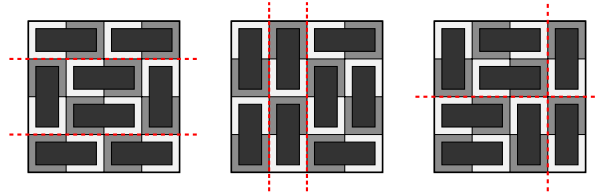


Figure 4: Some tilings have horizontal cuts, some have vertical cuts and some have both.

As another example consider the board B consisting of two 4×4 boards joint by two extra squares as shown in Figure 5.

In a tiling of B it is impossible that one extra square is covered by a domino containing a square from the left side, while the other extra square is covered by a domino containing a square from the right side: If we place the two dominoes as shown on the right in Figure 5, we are left with two disconnected parts of size 15. And since 15 is odd, this cannot work out. Hence the set \mathcal{T} of all tilings

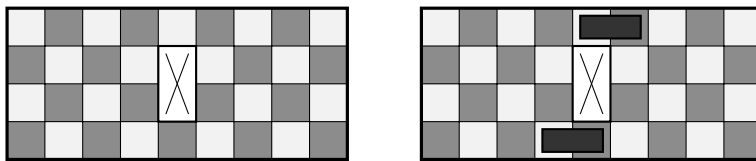


Figure 5: The board B consisting of two 4×4 boards and two extra squares connecting them as shown. The partial covering on the right cannot be extended to a tiling.

of B can be partitioned into

$$\begin{aligned}\mathcal{T}_1 &= \{T \mid T \text{ matches both extra squares to the left}\} \\ \mathcal{T}_2 &= \{T \mid T \text{ matches both extra squares to the right}\}.\end{aligned}$$

Meta Structure

We say two tilings are *adjacent*, if one can be transformed into the other by taking two dominoes lying like this $\begin{smallmatrix} \blacksquare & \blacksquare \end{smallmatrix}$ and turning them by 90 degrees so they are lying like this $\begin{smallmatrix} \blacksquare & \blacksquare \\ \blacksquare & \blacksquare \end{smallmatrix}$ (or vice versa). Call this a *turn operation*. If we draw all tilings of the 4×4 board and then draw a line between tilings that are adjacent, we get the picture on the left of Figure 6.

With this in mind, we can speak about the *distance* of two tilings, the number of turn operations required to transform one into the other. We could also look for a pair of tilings with maximum distance (this would be an optimization problem).

We can even discover a deeper structure, but for this we need to identify different types of turn operations. We can turn $\begin{smallmatrix} \blacksquare & \blacksquare \end{smallmatrix}$ into $\begin{smallmatrix} \blacksquare & \blacksquare \\ \blacksquare & \blacksquare \end{smallmatrix}$, call this an *up-turn*, or $\begin{smallmatrix} \blacksquare & \blacksquare \\ \blacksquare & \blacksquare \end{smallmatrix}$ into $\begin{smallmatrix} \blacksquare & \blacksquare \end{smallmatrix}$, call this a *side-turn*. The turn can happen on the background $\begin{smallmatrix} \square & \square \\ \square & \square \end{smallmatrix}$ or the background $\begin{smallmatrix} \square & \square \\ \square & \square \end{smallmatrix}$ (white squares form a falling or rising diagonal). We call an operation a *flip* if it is an up-turn on $\begin{smallmatrix} \square & \square \\ \square & \square \end{smallmatrix}$ or a side-turn on $\begin{smallmatrix} \square & \square \\ \square & \square \end{smallmatrix}$. Call it a *flop* otherwise.

Quite surprisingly, walking upward in the graph in Figure 6 always corresponds to flips and walking downwards corresponds to flops. This means that there is a natural partial ordering on the tilings: We can say a tiling B is *greater* than a tiling A if we can get from A to B by a sequence of flips. As it turns out, the order this gives has special properties. It is a so-called distributive lattice, in particular, there is a greatest tiling.

Optimization

Different tilings have a different set of *decreasing free paths*. Such a path (see Figure 7) proceeds monotonously from the top left corner of the board along the borders of squares to the bottom right, and does not “cut through” any domino.

Some tilings admit more such paths than others. It is conjectured that in an $m \times n$ board the maximum number is always attained by one of the two tilings where all dominoes are oriented the same way (see right of Figure 7). But no proof has been found yet.

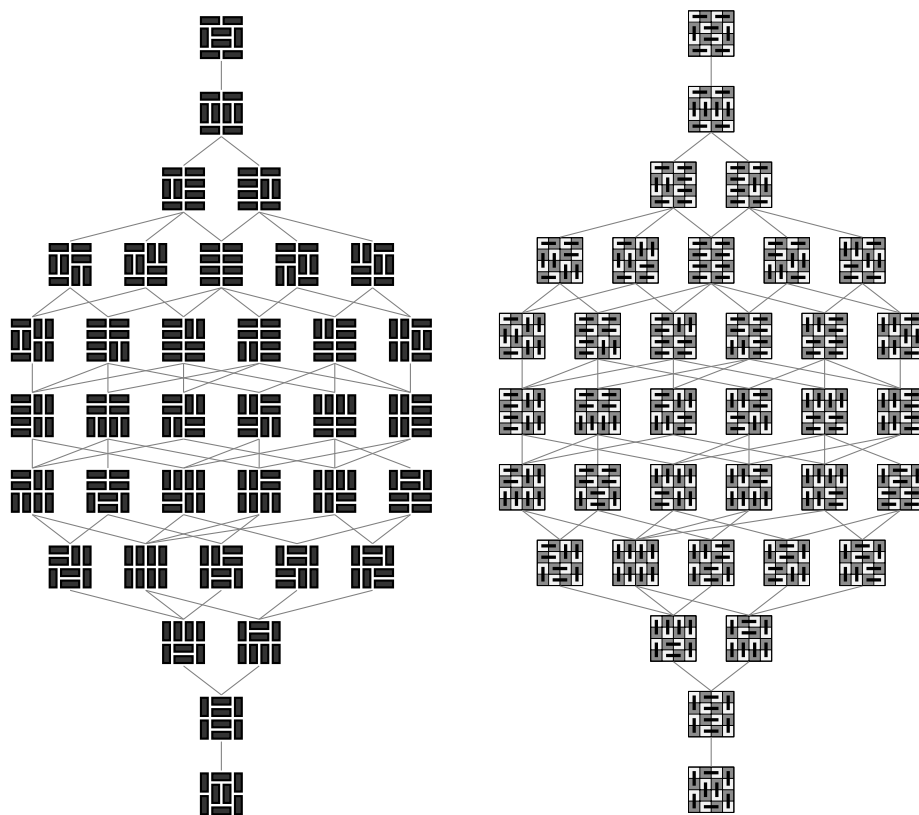


Figure 6: On the left: The set of all tilings of the 4×4 board. Two tilings are connected by an edge if they can be transformed into one another by a single flip. On the right: The same picture, but with the chessboard still drawn beneath the tilings, so you can check that upward edges correspond to flips (as defined in the text).

1 Permutations and Combinations

Some people mockingly say that combinatorics is merely about counting things. In the section at hand we will do little to dispel this prejudice.

We assume the reader is familiar with basic set theory and notions such as unions, intersections, Cartesian products and differences of two finite sets.

1.1 Basic Counting Principles

We list a few very simple and intuitive observations. Many of them were already used in the introduction. You may think that they are so easy that they do not even deserve a name. Still: If we take counting seriously, we should be aware of simple facts we implicitly use all the time.

1.1.3 Subtraction Principle

Let S be a subset of a finite set T . We define $\bar{S} := T \setminus S$, the complement of S in T . Then the *subtraction principle* claims that

$$|S| = |T| - |\bar{S}|.$$

Example. If T is the set of students studying at KIT and S the set of students studying neither math nor computer science. If we know $|T| = 23905$ and $|S| = 20178$, then we can compute the number $|\bar{S}|$ of students studying either math or computer science:

$$|\bar{S}| = |T| - |S| = 23905 - 20178 = 3727.$$

1.1.4 Bijection Principle

If S and T are finite sets, then the *bijection principle* claims that

$$|S| = |T| \quad \Leftrightarrow \quad \text{there exists a bijection between } S \text{ and } T.$$

Example. Let S be the set of students attending the lecture and T the set of homework submissions for the first problem sheet.

If the number of students and the number of submissions coincide, then there is a bijection between students and submissions¹ and vice versa.

We now consider two other principles that are similarly intuitive and natural but will frequently and explicitly occur as patterns in proofs.

1.1.5 Pigeonhole Principle

Let S_1, \dots, S_m be finite sets that are pairwise disjoint and $|S_1| + |S_2| + \dots + |S_m| = n$. Then the *pigeonhole principle* claims that

$$\exists i \in \{1, \dots, m\} : |S_i| \geq \left\lceil \frac{n}{m} \right\rceil \quad \text{and} \quad \exists j \in \{1, \dots, m\} : |S_j| \leq \left\lfloor \frac{n}{m} \right\rfloor.$$

Example. Assume there are 5 holes in the wall where pigeons nest. Say there is a set S_i of pigeons nesting in hole i . Assume there are $n = 17$ pigeons in total. Then we know:

- There is some hole with at least $\lceil 17/5 \rceil = 4$ pigeons.
- There is some hole with at most $\lfloor 17/5 \rfloor = 3$ pigeons.

1.1.6 Double counting

If we count the same quantity in two different ways, then this gives us a (perhaps non-trivial) identity. This method often helps get control on overcounting. Let A and B be finite sets and consider a set $S \subseteq A \times B$. Then the method of *double counting* claims that

$$\sum_{a \in A} |\{b \in B \mid (a, b) \in S\}| = |S| = \sum_{b \in B} |\{a \in A \mid (a, b) \in S\}|.$$

¹Whether a “natural” correspondence can easily be determined is an entirely different matter: That would require that you *cleanly* write *your name* onto your submission.

Example (Handshaking Lemma). Assume there are n people at a party and everybody will shake hands with everybody else. What is the total number N of handshakes that occur? If we just sum up for each guest each of its handshakes, then we overcount each handshake. In order to control this overcount we shall instead count the number M of pairs (P, S) where P is a guest and S is a handshake involving P . We count this number in two ways:

First way: There are n guests and everybody shakes $n - 1$ hands. So

$$M = \sum_{P \text{ guest}} |\{(P, S) \mid S \text{ handshake with } P\}| = \sum_{P \text{ guest}} (n - 1) = n \cdot (n - 1).$$

Second way: Every handshake involves two guests. So

$$M = \sum_{S \text{ handshake}} |\{(P, S) \mid P \text{ involved in } S\}| = \sum_{S \text{ handshake}} 2 = 2 \cdot N.$$

Together, we see that $n(n - 1) = M = 2N$ and hence we obtain the identity

$$N = \frac{n \cdot (n - 1)}{2}.$$

Example. In the example above we have seen a first way how to count the number N (already utilizing double counting). Here we count N in a second way. Label the guests from 1 to n . To avoid counting a handshake twice, we count for guest i only the handshakes with guests of smaller numbers. Then the total number of handshakes is

$$\sum_{i=1}^n (i - 1) = \sum_{i=0}^{n-1} i = \sum_{i=1}^{n-1} i.$$

Hence we obtain another identity

$$\sum_{i=1}^{n-1} i = N = \frac{n \cdot (n - 1)}{2}.$$

1.2 Ordered Arrangements – Strings, Maps and Products

Define $[n] := \{1, \dots, n\}$ to be the set of the first n natural number and let X be a finite set.

An *ordered arrangement* of n elements of X is a map $s : [n] \rightarrow X$. Ordered arrangements are so common that they occur in many situations and are known under different names. We might be inclined to say that “**Banana**” is a string (or word) and that “0815422372” is a sequence even though they are both ordered arrangements of characters. Depending on the circumstances different notation may be in order as well.

We will introduce the most common perspectives on what an ordered arrangement is and the accompanying names and notation. Throughout, we use the example $X = \{\square, \circ, \mid, \Delta, \nabla\}$ and the ordered arrangement s of $n = 7$ elements given as the table:

i	1	2	3	4	5	6	7
$s(i)$	\square	\circ	\circ	\triangle	\circ	∇	∇

function: We call $[n]$ the *domain* of s and $s(i)$ the *image* of i under s ($i \in [n]$). The set $\{x \in X \mid s(i) = x \text{ for some } i \in [n]\}$ is the *range* of s .

In the example, the domain of s is $\{1, 2, 3, 4, 5, 6, 7\}$, the image of 3 is \circ and the range of s is $\{\square, \circ, \triangle, \nabla\}$. Note that 1 is not in the range of s .

string: We call s an X -*string* of *length* n and write $s = s(1)s(2)s(3)\cdots s(n)$. The i -th position (or *character*) in s is denoted by $s_i = s(i)$. The set X is an *alphabet* and its elements are *letters*. Often s is called a *word*.

In the example, we would say that $s = \square\circ\circ\triangle\circ\nabla\nabla$ is a string (or word) of length n over the five-letter alphabet X . The fourth character of s is $s_4 = \triangle$.

tuple: We can view s as an element of the n -fold *Cartesian product* $X_1 \times \cdots \times X_n$, where $X_i = X$ for $i \in [n]$. We call s a *tuple* and write it as (s_1, s_2, \dots, s_n) . The element s_i is called the i -th *coordinate* ($i \in [n]$). Viewing arrangements as elements of products makes it easy to restrict the number of allowed values for a particular coordinate (just choose $X_i \subsetneq X$). In the example we would write $s = (\square, \circ, \circ, \triangle, \circ, \nabla, \nabla)$. Its first coordinate is \square .

sequence: If X is a set of numbers, then s is often called a *sequence* and s_i is the i -th *position*, *term* or *element* in the sequence. We would not use this terminology in the example above but could have, for instance, $s(i) := 3i + 2/5$. In that case the fourth term is equal to 12.4.

In the following we will mostly view arrangements as functions but will freely switch perspective when appropriate.

1.2.1 Permutations

The most important ordered arrangements are those in which the mapping is injective, i.e., $s(i) \neq s(j)$ for $i \neq j$.

Definition 1.1 (Permutation). Let X be a finite set. We define

permutation: A *permutation* of X is a bijective map $\pi : [n] \rightarrow X$. Usually we choose $X = [n]$ and denote the set of all permutations of $[n]$ by S_n .

We tend to write permutations as strings if $n < 10$, take for example $\pi = 2713546$ by which we mean the function:

i	1	2	3	4	5	6	7
$\pi(i)$	2	7	1	3	5	4	6

k -Permutation: For $1 \leq k \leq |X|$ a k -*permutation* of X is an ordered arrangement of k distinct elements of X , i.e. an injective map $\pi : [k] \rightarrow X$. The set of all k -permutations of $X = [n]$ is denoted by $P(n, k)$. In particular we have $S_n = P(n, n)$.

Circular k -Permutation: We say that two k -permutations $\pi_1, \pi_2 \in P(n, k)$ are *circular equivalents* if there exists a shift $s \in [k]$ such that the following implication holds:

$$i + s \equiv j \pmod{k} \Rightarrow \pi_1(i) = \pi_2(j)$$

This equivalence relation partitions $P(n, k)$ into equivalence classes. A class is called a *circular k -permutation*. The set of all circular k -permutations is denoted by $P_c(n, k)$.

Take for example $\pi_1 = 76123$, $\pi_2 = 12376$ and $\pi_3 = 32167$. Then π_1 and π_2

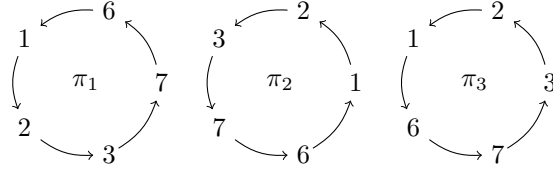


Figure 8: Visualization of the 5-permutations π_1 , π_2 and π_3 where values are arranged in a circular fashion (the first value being on the right of the cycle). Note that π_1 and π_2 can be transformed into one another by “turning”. They are therefore circular equivalents. Flipping is not allowed however, so π_3 is not equivalent to π_1 and π_2 .

are circular equivalents as witnessed by the shift $s = 3$. They are therefore representatives of the same circular 5-permutation (with elements from $[7]$). π_3 belongs to a different circular 5-permutation. Consider Figure 8 for a visualization.

We now count the number of (circular) k -permutations. For this we need the notation of factorials: $n! := 1 \cdot 2 \cdot \dots \cdot n$.

Theorem 1.2. *For any natural numbers $1 \leq k \leq n$ we have*

$$(i) |P(n, k)| = n \cdot (n - 1) \cdot \dots \cdot (n - k + 1) = \frac{n!}{(n-k)!}.$$

$$(ii) |P_c(n, k)| = \frac{n!}{k \cdot (n-k)!}.$$

Proof. (i) A permutation is an injective map $\pi : [k] \rightarrow [n]$. We count the number of ways to pick such a map, picking the images one after the other. There are n ways to choose $\pi(1)$. Given a value for $\pi(1)$, there are $n - 1$ ways to choose $\pi(2)$ (we may not choose $\pi(1)$ again). Continuing like this, there are $n - i + 1$ ways to pick $\pi(i)$ and the last value we pick is $\pi(k)$ with $n - k + 1$ possibilities. Every k -permutation can be constructed like this in exactly one way.

The total number of k -permutations is therefore given as the product:

$$|P(n, k)| = n \cdot (n - 1) \cdot \dots \cdot (n - k + 1) = \frac{n!}{(n - k)!}.$$

(ii) We count $P(n, k)$ in two ways:

First way: $|P(n, k)| = \frac{n!}{(n-k)!}$ which we proved in (i).

Second way: $|P(n, k)| = |P_c(n, k)| \cdot k$ because every equivalence class in $P_c(n, k)$ contains k permutations from $P(n, k)$ (since there are k ways to rotate a k -permutation).

From this we get $\frac{n!}{(n-k)!} = |P_c(n, k)| \cdot k$ which implies the claim. \square

1.3 Unordered Arrangements – Combinations, Subsets and Multisets

Let X be a finite set and $k \in \mathbb{N}$. An *unordered arrangement* of k elements of X is a multiset S of size k with elements from X .

Take for example $X = \{\square, \circ, \triangle, \nabla\}$, then an unordered arrangement of 7 elements could be $S = \{\square, \square, \circ, \triangle, \triangle, \triangle, \nabla\}$. Order in sets and multisets does not matter, so we could write the same multiset as $S = \{\triangle, \square, \triangle, \nabla, \square, \triangle, \circ\}$.

However, we prefer the following notation, where every *type* $x \in X$ occurring in S is given only once and accompanied by its *repetition number* r_x that captures how often the type occurs in the multiset. The example above is written as: $S = \{2 \cdot \square, 1 \cdot \circ, 3 \cdot \triangle, 1 \cdot \nabla\}$ and we would say the type \square has repetition number 2, the type \circ has repetition number 1 and so on. We write $\square \in S$, $r_{\square} = 2$, $\circ \in S$, $r_{\circ} = 1$ and so on.

The difference between ordered and unordered arrangements is that ordered arrangements are selections of elements of X that are done one at a time, while unordered arrangements are selections of objects done all at the same time.

A typical example for unordered arrangements would be shopping lists: You may need three bananas and two pears, but this is the same as needing a pear, three bananas and another pear. In a sense, you need everything *at once* with no temporal ordering. This is in contrast to something like a telephone number (which is an ordered arrangement) where dialing a 5 first and a 9 later is different from dialing 9 first and 5 later.

As with ordered arrangements, the most important case for unordered arrangements is that all repetition numbers are 1, i.e. $r_x = 1$ for all $x \in S$. Then S is simply a *subset* of X , denoted by $S \subseteq X$.

Definition 1.3. k -Combination: Let X be a finite set. A k -combination of X is an unordered arrangement of k distinct elements from X . We prefer the more standard term *subset* and use “combination” only when we want to emphasize the selection process. The set of all k -subsets of X is denoted by $\binom{X}{k}$ and if $|X| = n$ then we denote

$$\binom{n}{k} := \left| \binom{X}{k} \right|.$$

k -Combination of a Multiset: Let X be a finite set of types and let M be a finite multiset with types in X and repetition numbers $r_1, \dots, r_{|X|}$. A k -combination of M is a multiset with types in X and repetition numbers $s_1, \dots, s_{|X|}$ such that $s_i \leq r_i$, $1 \leq i \leq |X|$, and $\sum_{i=1}^{|X|} s_i = k$.

If for example $M = \{2 \cdot \square, 1 \cdot \circ, 3 \cdot \triangle, 1 \cdot \nabla\}$, then $T = \{1 \cdot \square, 2 \cdot \triangle\}$ is a 3-combination of the multiset M , but $T' = \{3 \cdot \square, 0 \cdot \nabla\}$ is not.

k -Permutation of a Multiset: Let M be a finite multiset with set of types X .

A k -permutation of M is an ordered arrangement of k elements of M where different orderings of elements of the same type are not distinguished. This is a an ordered multiset with types in X and repetition numbers $s_1, \dots, s_{|X|}$ such that $s_i \leq r_i$, $1 \leq i \leq |X|$, and $\sum_{i=1}^{|X|} s_i = k$.

Note that there might be several elements of the same type compared to a permutation of a set (where each repetition number equals 1). If for example $M = \{2 \cdot \square, 1 \cdot \circ, 3 \cdot \Delta, 1 \cdot \nabla\}$, then $T = (\square, \Delta, \Delta, \square)$ is a 4-permutation of the multiset M .

Theorem 1.4. For $0 \leq k \leq n$, we have

$$P(n, k) = \binom{n}{k} \cdot k! \quad \text{and therefore} \quad \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Proof. To build an element from $P(n, k)$, we first choose k elements from $[n]$, by our definition of $\binom{n}{k}$, there are exactly $\binom{n}{k}$ ways to do so. Then choose an order of the k elements, there are $|P(k, k)| = \frac{k!}{0!} = k!$ ways to do so.

Every element from $P(n, k)$ can be constructed like this in exactly one way so $|P(n, k)| = \binom{n}{k} \cdot k!$ which proves the claim.

Using Theorem 1.2(i) now gives the identity on the right as well. \square

The numbers $\binom{n}{k}$ for $0 \leq k \leq n$ are called *binomial coefficients* because these are the coefficients when expanding the n -th power of a sum of two variables:

$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}.$$

1.4 Multinomial Coefficients

The binomial formula can be generalized for more than two variables.

Definition 1.5. For non-negative integers k_1, \dots, k_r with $k_1 + \dots + k_r = n$ the *multinomial coefficients* $\binom{n}{k_1, \dots, k_r}$ are the coefficients when expanding the n -th power of a sum of r variables. In other words, define them to be the unique numbers satisfying the following identity:

$$(x_1 + \dots + x_r)^n = \sum_{\substack{k_1, \dots, k_r \\ k_1 + \dots + k_r = n}} \binom{n}{k_1, \dots, k_r} x_1^{k_1} x_2^{k_2} \dots x_r^{k_r}.$$

Example. Verify by multiplying out:

$$\begin{aligned} (x_1 + x_2 + x_3)^4 &= 1 \cdot (x_1^4 + x_2^4 + x_3^4) \\ &+ 4 \cdot (x_1^3 x_2 + x_1^3 x_3 + x_1 x_2^3 + x_2^3 x_3 + x_1 x_3^3 + x_2 x_3^3) \\ &+ 6 \cdot (x_1^2 x_2^2 + x_1^2 x_3^2 + x_2^2 x_3^2) \\ &+ 12 \cdot (x_1^2 x_2 x_3 + x_1 x_2^2 x_3 + x_1 x_2 x_3^2). \end{aligned}$$

The coefficient in front of $x_1^2 x_2 x_3$ is 12, the coefficient in front of $x_1^2 x_3^2$ is 6 and the coefficient in front of x_3^4 is 1.

According to the last definition this means:

$$\binom{4}{2,1,1} = 12, \quad \binom{4}{2,0,2} = 6, \quad \binom{4}{0,0,4} = 1.$$

Theorem 1.6 (Multinomial Theorem).

For non-negative integers k_1, \dots, k_r with $k_1 + \dots + k_r = n$ we have:

$$\binom{n}{k_1, \dots, k_r} = \binom{n}{k_1} \cdot \binom{n-k_1}{k_2} \cdot \dots \cdot \binom{n-k_1-\dots-k_{r-1}}{k_r} = \frac{n!}{k_1! \cdot k_2! \cdot \dots \cdot k_r!}.$$

Proof. The naive way to multiply out $(x_1 + \dots + x_r)^n$ would be:

$$(x_1 + \dots + x_r)^n = \underbrace{\sum_{i_1=1}^r \sum_{i_2=1}^r \dots \sum_{i_n=1}^r}_{n \text{ summation signs}} x_{i_1} x_{i_2} \dots x_{i_n}.$$

The monomial $x_{i_1} x_{i_2} \dots x_{i_n}$ is equal to $x_1^{k_1} \dots x_r^{k_r}$ if from the indices $\{i_1, \dots, i_n\}$ exactly k_1 are equal to 1, k_2 equal to 2 and so on.

We count the number of assignments of values to the indices $\{i_1, \dots, i_n\}$ satisfying this.

Choose k_1 indices to be equal to 1: There are $\binom{n}{k_1}$ ways to do so.

Choose k_2 indices to be equal to 2: There are $n - k_1$ indices left to choose from, so there are $\binom{n-k_1}{k_2}$ ways to choose k_2 of them.

Choose k_j indices to be equal to j ($j \in [r]$): There $n - k_1 - \dots - k_{j-1}$ indices left to choose from, so there are $\binom{n-k_1-\dots-k_{j-1}}{k_j}$ ways to choose k_j of them.

Hence the multinomial coefficient (i.e. the coefficient of $x_1^{k_1} x_2^{k_2} \dots x_r^{k_r}$) is

$$\binom{n}{k_1, \dots, k_r} = \binom{n}{k_1} \cdot \binom{n-k_1}{k_2} \cdot \dots \cdot \binom{n-k_1-k_2-\dots-k_{r-1}}{k_r}$$

and the first identity is proved. Now use Theorem 1.4 to rewrite the binomial coefficients and obtain

$$\frac{n!}{k_1!(n-k_1)!} \cdot \frac{(n-k_1)!}{k_2!(n-k_1-k_2)!} \cdot \dots \cdot \frac{(n-k_1-\dots-k_{r-1})!}{k_r!(n-k_1-\dots-k_r)!}$$

Conveniently, many of the factorial terms cancel out, like this:

$$\frac{n!}{k_1!(\cancel{n-k_1})!} \cdot \frac{(\cancel{n-k_1})!}{k_2!(\cancel{n-k_1-k_2})!} \cdot \dots \cdot \frac{(\cancel{n-k_1-\dots-k_{r-1}})!}{k_r!(n-k_1-\dots-k_r)!}.$$

Also, $(n - k_1 - \dots - k_r)! = (n - n)! = 0! = 1$ so we end up with:

$$\frac{n!}{k_1! \cdot k_2! \cdot \dots \cdot k_r!}.$$

□

Example. $\binom{5}{3,0,2} = \frac{5!}{3!0!2!} = \frac{5 \cdot 4 \cdot \cancel{3} \cdot \cancel{2} \cdot 1}{\cancel{3} \cdot \cancel{2} \cdot 1 \cdot 2 \cdot 1} = 10.$

Note that the last Theorem establishes that binomial coefficients are special cases of multinomial coefficients. We have for $0 \leq k \leq n$:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \binom{n}{k, n-k}.$$

We extend the definition of binomial and multinomial coefficients by setting $\binom{n}{k_1, \dots, k_r} = 0$ if $k_i = -1$ for some i , and $\binom{n}{-1} = \binom{n}{n+1} = 0$. This makes stating the following lemma more convenient.

Lemma 1.7 (Pascal's Formula).

If $n \geq 1$ and $0 \leq k \leq n$, we have

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}.$$

More generally, for $n \geq 1$ and $k_1, \dots, k_r \geq 0$ with $k_1 + \dots + k_r = n$, we have

$$\binom{n}{k_1, \dots, k_r} = \sum_{i=1}^r \binom{n-1}{k_1, \dots, k_i-1, \dots, k_r}.$$

Proof. Note first, that in the case of binomial coefficients, the claim can be rewritten as:

$$\binom{n}{k, n-k} = \binom{n-1}{k, n-k-1} + \binom{n-1}{k-1, n-k}$$

so it really is a special case of the second claim, which we prove now. By definition of multinomial coefficients we have the identity:

$$(x_1 + \dots + x_r)^n = \sum_{\substack{k_1, \dots, k_r \geq 0 \\ k_1 + \dots + k_r = n}} \binom{n}{k_1, \dots, k_r} x_1^{k_1} \dots x_r^{k_r}.$$

Exploring a different way to expand $(x_1 + \dots + x_r)^n$, we obtain:

$$\begin{aligned} (x_1 + \dots + x_r)^n &= (x_1 + \dots + x_r) \cdot (x_1 + \dots + x_r)^{n-1} \\ &= (x_1 + \dots + x_r) \cdot \sum_{\substack{k'_1, \dots, k'_r \geq 0 \\ k'_1 + \dots + k'_r = n-1}} \binom{n-1}{k'_1, \dots, k'_r} x_1^{k'_1} \dots x_r^{k'_r} \\ &= \sum_{i=1}^r \sum_{\substack{k'_1, \dots, k'_r \geq 0 \\ k'_1 + \dots + k'_r = n-1}} \binom{n-1}{k'_1, \dots, k'_r} x_1^{k'_1} \dots x_i^{k'_i+1} \dots x_r^{k'_r}. \end{aligned}$$

By substituting indices as $k_i := k'_i + 1$ and $k_j := k'_j$ (for $j \neq i$) this is equal to:

$$(x_1 + \dots + x_r)^n = \sum_{i=1}^r \sum_{\substack{k_1, \dots, k_r \geq 0 \\ k_i-1 \geq 0 \\ k_1 + \dots + k_r = n}} \binom{n-1}{k_1, \dots, k_i-1, \dots, k_r} x_1^{k_1} \dots x_r^{k_r}.$$

Note that the condition $k_i - 1 \geq 0$ is not needed, because for the summands with $k_i - 1 = -1$ we defined the coefficient under the sum to be 0. We remove it and swap the summation signs, ending up with:

$$(x_1 + \dots + x_r)^n = \sum_{\substack{k_1, \dots, k_r \geq 0 \\ k_1 + \dots + k_r = n}} \sum_{i=1}^r \binom{n-1}{k_1, \dots, k_i-1, \dots, k_r} x_1^{k_1} \dots x_r^{k_r}$$

Now we have two ways to write the coefficients of $(x_1 + \dots + x_r)^n$ and therefore the identity:

$$\binom{n}{k_1, \dots, k_r} = \sum_{i=1}^r \binom{n-1}{k_1, \dots, k_i-1, \dots, k_r}$$

as claimed. \square

You may already know Pascal's Triangle. It is a way to arrange binomial coefficients in the plane as shown in Figure 9. It is possible to do something similar in the case of multinomial coefficients, however, when drawing all coefficients of the form $\binom{n}{k_1, \dots, k_r}$, the drawing will be r -dimensional. For $r = 3$, a "Pascal Pyramid" is given in Figure 10.

Now that we know what multinomial coefficients are and how to compute them, it is time to see how they can help us count things.

Example. How many 6-permutations are there of the multiset $\{\circ, \Delta, \Delta, \square, \square, \square\}$? Trying to list them all $((\circ, \Delta, \Delta, \square, \square, \square), (\Delta, \circ, \Delta, \square, \square, \square), (\Delta, \Delta, \circ, \square, \square, \square), \dots)$ would be tedious. Well, there are 6 possible positions for \circ . Then there are 5 positions left, 2 of which need to contain Δ . The three \square need to go to the three remaining positions, so there is just one choice left for them. This means there are $6 \cdot \binom{5}{2} \cdot 1 = 60$ arrangements in total. This is equal to $\binom{6}{1} \binom{5}{2} \binom{3}{3} = \binom{6}{1,2,3}$, which is no coincidence:

Theorem 1.8. *Let S be a finite multiset with k different types and repetition numbers r_1, r_2, \dots, r_k . Let the size of S be $n = r_1 + r_2 + \dots + r_k$. Then the number of n -permutations of S equals*

$$\binom{n}{r_1, \dots, r_k}.$$

Proof. Label the k types as a_1, \dots, a_k . In an n -permutation there are n positions that need to be assigned a type. First choose the r_1 positions for the first type, there are $\binom{n}{r_1}$ ways to do so. Then assign r_2 positions for the second type, out of the $n - r_1$ positions that are still left to choose. That amounts to $\binom{n-r_1}{r_2}$ choices. Continuing like this, the total number of choices will be:

$$\binom{n}{r_1} \cdot \binom{n-r_1}{r_2} \cdot \dots \cdot \binom{n-r_1-r_2-\dots-r_{k-1}}{r_k} \stackrel{\text{Thm 1.6}}{=} \binom{n}{r_1, r_2, \dots, r_k}.$$

\square

We have seen before how the coefficient $\binom{n}{k} = \binom{n}{k, n-k}$ counts the number of k -combinations of $[n]$ (i.e. number of subsets of $[n]$). Now we learned, that

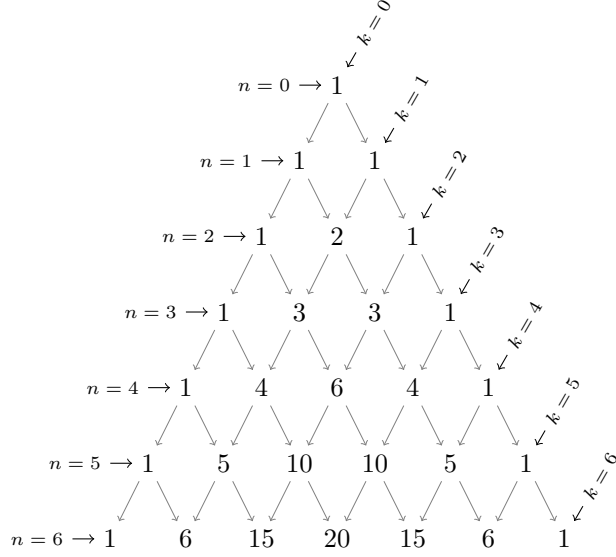
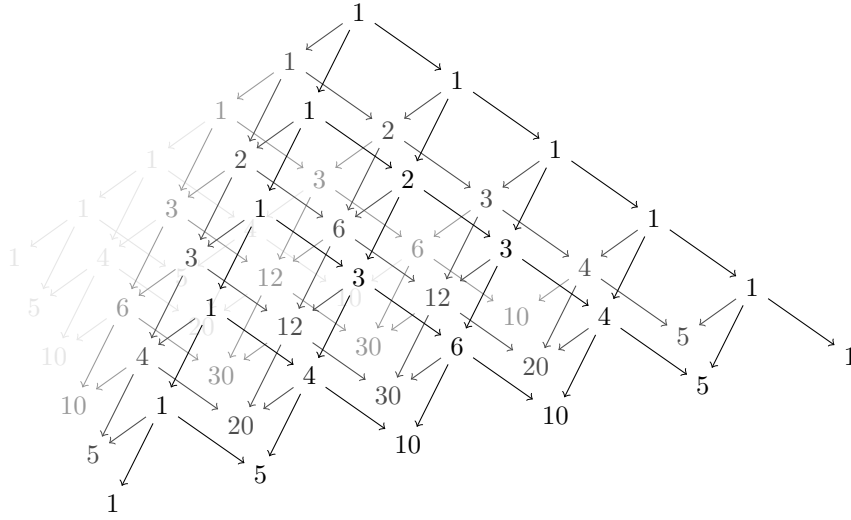


Figure 9: Arrangement of the binomial coefficients $\binom{n}{k}$. Lemma 1.7 shows that the number $\binom{n}{k}$ is obtained as the sum of the two numbers $\binom{n-1}{k-1}$ and $\binom{n-1}{k}$ directly above it.



it also counts the number of n -permutations of a multiset with two types and repetition number k and $n - k$. How come ordered and unordered things are counted by the same number? We demystify this by finding a natural bijection:

Consider the multiset $M := \{k \cdot \checkmark, (n - k) \cdot \times\}$ with types \checkmark and \times (the “chosen” type and the “unchosen” type). Now associate with an n -permutation of M the set of positions that contain \checkmark , for instance with $n = 5$, $k = 2$ and the permutation $(\checkmark, \times, \times, \checkmark, \times)$, the corresponding set would be $\{1, 4\} \subset [5]$ (since the first and fourth position received \checkmark). It is easy to see that every n -permutation of M corresponds to a unique k element subset of $[n]$ and vice versa.

Note that so far we only considered n -permutations of multisets of size n . What about general r -permutations of multisets of size n ? For example, the number of 2 permutations of the multiset $\{\square, \triangle, \triangle, \triangle, \circ\}$ is 7, since there is:

$$\square\triangle, \square\circ, \triangle\square, \triangle\triangle, \triangle\circ, \circ\square, \circ\triangle.$$

Note that $\circ\circ$ and $\square\square$ is not possible, since we have only of one copy of \circ and \square at our disposal. The weird number 7 already suggests that general r -permutations of n element multisets may not be as easy to count. Indeed, there is no simple formula as in Theorem 1.8 but we will see an answer using *the principle of inclusion and exclusion* later.

There is a special case other than $r = n$ that we can handle, though: If all repetition numbers r_i of a multiset with k types are bigger or equal than r , for instance when considering 2-permutations of $M := \{\square, \square, \triangle, \triangle, \triangle, \circ, \circ\}$, then those repetition numbers do not actually impose a restriction, since we will never run out of copies of any type. We would sometimes sloppily write $M = \{\infty \cdot \square, \infty \cdot \triangle, \infty \cdot \circ\}$ where the infinity sign indicates that there are “many” copies of the corresponding elements. The number of r -permutations of M is then equal to k^r , just choose one type of the k types for each of the r positions.

After permutations of multisets, we now consider combinations.

Example. Say you are told to bring two pieces of fruit from the supermarket and they got 🍌, 🍎 and 🍌 (large quantities of each). How many choices do you have? Well, there is: $\{\text{🍌}, \text{🍌}\}, \{\text{🍌}, \text{🍎}\}, \{\text{🍌}, \text{🍌}\}, \{\text{🍎}, \text{🍎}\}, \{\text{🍎}, \text{🍌}\}, \{\text{🍌}, \text{🍌}\}$, so six combinations. Note that bringing a 🍌 and an 🍎 is the same as bringing an 🍎 and a 🍌 (your selection is not ordered), so this option is counted only once.

We now determine the number of combinations for arbitrary number of types and number of elements to choose.

Theorem 1.9. *Let $r, k \in \mathbb{N}$ and let S be a multiset with k types and large repetition numbers (each r_1, \dots, r_k is at least r), then the number of r -combinations of S equals*

$$\binom{k + r - 1}{r}.$$

Proof. For clarity, we do the proof alongside an example. Let the types be a_1, a_2, \dots, a_k , for instance $k = 4$ and $a_1 = \text{🍌}$, $a_2 = \text{🍎}$, $a_3 = \text{🍌}$, $a_4 = \text{🍌}$. Then imagine the r -combinations laid-out linearly, first all elements of type a_1 then all of type a_2 and so on. In our example this could be

$$\text{🍌🍌🍌🍌🍌} \text{ for the combination } \{2 \cdot \text{🍌}, 2 \cdot \text{🍌}, 1 \cdot \text{🍌}\}.$$

Now for each $i \in [k-1]$, draw a delimiter between types a_i and a_{i+1} , in our example:



Note that, since we have no elements of type $a_2 = \text{orange}$, there are two delimiters directly after one another. Given these delimiters, drawing the elements is actually redundant, just replacing every element with “•” yields:

• • | | • • | •.

The original k -combination can be reconstructed from this sequence of | and •: Just replace every • with a_1 until the first occurrence of |, then use a_2 until the next |, then a_3 and so on. So every $(r+k-1)$ -permutation of $T := \{(k-1) \cdot |, r \cdot \bullet\}$ corresponds to an r -combination of S and vice versa.

We know how to count the former, by Theorem 1.8 the number of $(r+k-1)$ -permutations of T is $\binom{r+k-1}{r, k-1} = \binom{r+k-1}{r}$. \square

Counting r -combinations of multisets where repetition numbers r_1, \dots, r_k may be smaller than r is more difficult. We will see later how the *inclusion-exclusion principle* provides an answer in this case.

1.5 The Twelfold Way – Balls in Boxes

The most classic combinatorial problem concerns counting arrangements of n balls in k boxes. There are four cases: $\mathbf{U} \rightarrow \mathbf{L}$, $\mathbf{L} \rightarrow \mathbf{U}$, $\mathbf{L} \rightarrow \mathbf{L}$, $\mathbf{U} \rightarrow \mathbf{U}$, for arrangements of **L**abeled or **U**nabeled balls in **L**abeled or **U**nabeled boxes. Here “labeled” means distinguishable and “unlabeled” means indistinguishable.

These four cases become twelve cases if we consider the following sub-variants:

- (i) No box may contain more than one ball. Example: When assigning students to topics in a seminar, there may be at most one student per topic.
- (ii) Each box must contain at least one ball. Example: When you want to get 10 people and 5 cars to Berlin, you have some flexibility in distributing the people to cars, but a car cannot drive on its own.
- (iii) No restriction.

A summary of the results of this section is given in the end in Table 1.

Instead of counting the ways balls can be arranged in boxes, some people count ways that balls can be *put into* the boxes or being *picked from* the boxes, but this does not actually make a difference. We now systematically examine all 12 cases.

1.5.1 $\mathbf{U} \rightarrow \mathbf{L}$: n Unlabeled Balls in k Labeled Boxes

Example 1.10. Torsten decided there should always be $k = 30$ points to a worksheet and $n = 5$ problems per worksheet. Stefan believes some problems are easier than others and deserve more points. He wonders how many ways there are to distribute the points to the problems. In this setting, balls correspond

to points (points are not labeled, they are just points) and boxes correspond to problems (they are labeled: There is problem 1,2,3,4 and 5).

In other words, we search for solutions to the equation

$$30 = x_1 + x_2 + x_3 + x_4 + x_5$$

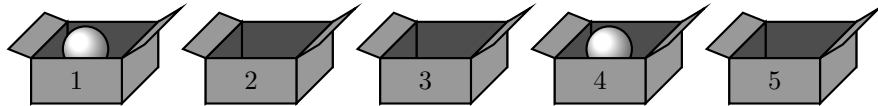
with non-negative integers x_1, \dots, x_5 , for example:

$$30 = 8 + 4 + 5 + 5 + 7 \quad \text{or} \quad 30 = 10 + 5 + 8 + 4 + 3.$$

The students say that they would like it if some of the problems are “bonus problems” worth zero points, so the partition $30 = 10 + 10 + 10 + 0 + 0$ should be permissible. Torsten remains unconvinced.

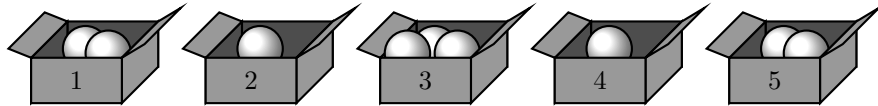
We come back to this later and examine three cases for general n and k in the balls-and-boxes formulation:

≤ 1 ball per box Of course, this is only possible if there are at most as many balls as boxes ($n \leq k$). For $n = 2$ and $k = 5$ one arrangement would be:



Each of the k boxes can have two states: Occupied (one ball in it) or empty (no ball in it) and exactly n boxes are occupied. The number of ways to choose these occupied boxes is $\binom{k}{n}$.

≥ 1 ball per box Of course, this is only possible if there are at least as many balls as boxes ($n \geq k$). For example for $n = 9$ and $k = 5$:



To count the number of ways to do this, arrange the balls linearly, like this:



and choose $k - 1$ out of $n - 1$ gaps between the balls that correspond to the beginning of a new box. In the example above this would look like this:



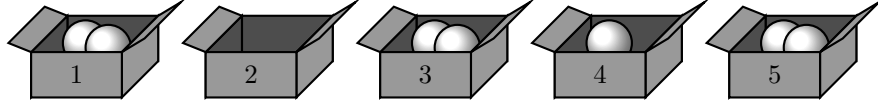
There is a bijection between the arrangements with no empty box and the choices of $n - 1$ gaps for delimiters out of $k - 1$ gaps in total. We know how to count the latter: There are $\binom{n-1}{k-1}$ possibilities.

Going back to the example, we now know there are $\binom{29}{5}$ solutions to the equation:

$$x_1 + x_2 + x_3 + x_4 + x_5 = 30$$

where x_1, \dots, x_5 are *positive* integers.

arbitrary number of balls per box Now boxes are allowed to contain any number of elements, including 0. One example for $n = 7$ and $k = 5$ would be:



We count the number of such arrangements in three different ways:

1. From an arrangement we obtain a string by “reading” it from left to right, writing \bullet when we see a ball and $|$ when a new box starts. For the example above this yields the string



We see: The permutations of the multiset $\{n \cdot \bullet, (k-1) \cdot | \}$ correspond directly to the arrangements of the balls. Note the difference to the case of non-empty boxes we discussed before.

We know how to count the permutations of the multiset, by Theorem 1.8 there are $\binom{n+k-1}{k-1}$ of them.

2. There is a bijection between:
 - (i) Arrangements of n balls in k boxes
 - (ii) Arrangements of $n+k$ balls in k boxes with no empty box.

For a map from (i) to (ii), just add one ball to each box. For the inverse map from (ii) to (i) just remove one ball from each box.

We already know how to count (ii), there are $\binom{n+k-1}{k-1}$ such arrangements.

3. Assume $n, k \geq 1$. To count the arrangements, first choose the number i of boxes that should be empty ($0 \leq i \leq k-1$), then choose which boxes that should be (there are $\binom{k}{i}$ choices) and then distribute all n balls to the remaining boxes $k-i$ boxes such that non of those boxes is empty (we already know how to count this). This yields:

$$\sum_{i=0}^{k-1} \binom{k}{i} \binom{n-1}{k-i-1}.$$

This looks different from the other two results, which means we “accidentally” proved the non-trivial identity:

$$\binom{n+k-1}{k-1} = \sum_{i=0}^{k-1} \binom{k}{i} \binom{n-1}{k-i-1} \quad (n, k \geq 1).$$

Going back to the example, we now know there are $\binom{30+6-1}{6-1}$ solutions to the equation:

$$30 = x_1 + x_2 + x_3 + x_4 + x_5$$

with non-negative integers, i.e. that many assignments of points to the five exercises on exercise sheets such that the total number of points is 30.

Torsten thinks that no exercise should be worth more than 10 points. In the balls and boxes setting this limits the capacity r_i of a box i . This makes counting much more difficult but we will see a way to address this in Chapter 2 using the principle of inclusion and exclusion.

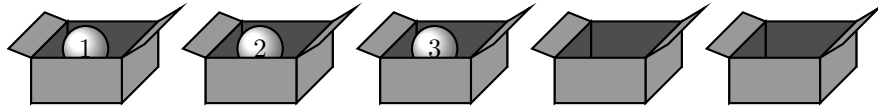
(As for homework problems, it turns out that Jonathan put `\def\points{6}` into his latex preamble, which settles the issue.)

1.5.2 $L \rightarrow U$: n Labeled Balls in k Unlabeled Boxes

Example 1.11. There are $n = 25$ kids on a soccer field that want to form $k = 4$ teams. The kids are on different skill levels (e.g. a team of five bad players is quite a different thing than a team of five good players) but the teams are unlabeled (it wouldn't make a difference if you swap all players of two teams). In how many ways can the n kids form k teams? In other words: In how many ways can the set of kids be *partitioned* into k parts?

Here, kids correspond to balls and teams correspond to boxes. Again, we consider three subcases for general k and n .

≤ 1 ball per box Of course, this is only possible if there are at most as many balls as boxes ($n \leq k$). Each ball will be in its own box. If the balls are labeled with numbers from $[n]$, there will be one box with the ball with label 1, one box with the ball with label 2 and so on. In other words, there is only 1 possibility, for $n = 3$ and $k = 5$ the only arrangement looks like this:



≥ 1 ball per box Of course, this is only possible if there are at least as many balls as boxes ($n \geq k$).

This is the same as the number of partitions of $[n]$ into k non-empty parts which we also call *Stirling Numbers of the second kind* and write as $s_k^{II}(n)$.

Some values are easy to determine:

- $s_0^{II}(0) = 1$: There is one way to partition the empty set into non-empty parts: $\emptyset = \bigcup_{X \in \emptyset} X$. Each $X \in \emptyset$ is non-empty (because no such X exists).
- $s_0^{II}(n) = 0$ (for $n \geq 1$): There is no way to partition non-empty sets into zero parts.
- $s_1^{II}(n) = 1$ (for $n \geq 1$): Every non-empty set X can be partitioned into one non-empty set in exactly one way: $X = X$.
- $s_2^{II}(n) = \frac{2^n - 2}{2} = 2^{n-1} - 1$ (for $n \geq 1$): We want to partition $[n]$ into two non-empty parts. If we consider the parts labeled (there is a first part and a second part), then choosing the first part fully determines the second and vice versa. Every subset of $[n]$ is allowed to be the first part – except for \emptyset and $[n]$. This amounts to $2^n - 2$ possibilities, however, since the parts are actually unlabeled (there is no “first” or “second”) every possibility is counted twice so we need to divide by 2.

- $s_n^{II}(n) = 1$: There is only one way to partition $[n]$ into n parts: Every number gets its own part.

A recursion formula for $s_k^{II}(n)$ is given as:

$$s_k^{II}(n) = k s_k^{II}(n-1) + s_{k-1}^{II}(n-1).$$

To see this, count the arrangements of the n labeled balls in unlabeled boxes with no empty box as follows:

- The ball of label n may have its own box (with no other ball in it). The number of such arrangements is equal to the number of arrangements of the remaining $n-1$ balls in $k-1$ boxes such that none of those $k-1$ boxes is empty. There are $s_{k-1}^{II}(n-1)$ of those.
- The box with the ball of label n contains another ball. Then, when removing ball n , there is still at least one ball per box. So removing ball n gets us to an arrangements of $n-1$ balls in k non-empty boxes. There are $s_k^{II}(n-1)$ of those and for each there are k possibilities where ball n could have been before removal (note that the boxes are distinguished by the balls that are already in it). So there are $k \cdot s_k^{II}(n-1)$ arrangements where ball n is not alone in a box.

The recursion formula follows from summing up the values for the two cases.

Note that the Stirling Numbers fulfill a recursion similar to the recursion of binomial coefficients, so there is something similar to the Pascal Triangle. See Figure 11.

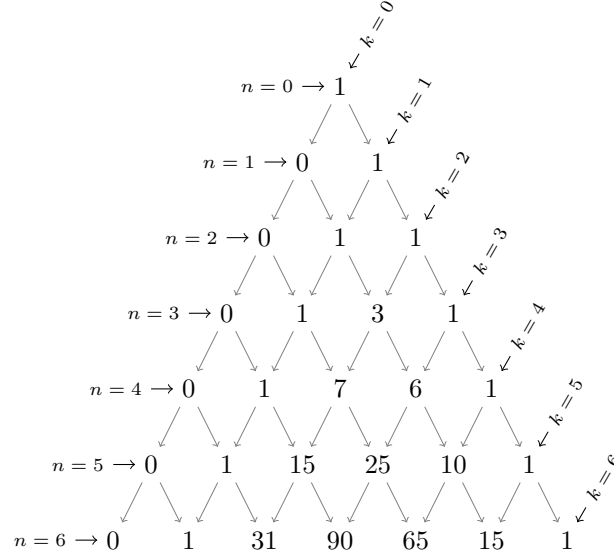


Figure 11: “Stirling Triangle”. A number $s_k^{II}(n)$ is obtained as the sum of the number $s_{k-1}^{II}(n-1)$ toward the top left and k times the number $s_k^{II}(n-1)$ towards the top right. E.g. for $n=6, k=3$: $90 = 15 + 3 \cdot 25$.

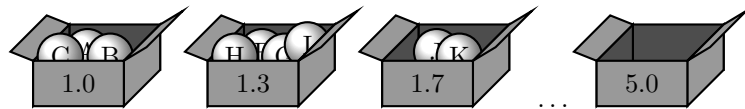
We will later see a closed form of $s_k^{II}(n)$ in Chapter 2 using the principle of inclusion and exclusion.

arbitrary number of balls per box Empty boxes are allowed. To count all arrangements, first choose the number i of boxes that should be non-empty ($0 \leq i \leq k$), then count the number of arrangements of n balls into the i boxes such that non of them is empty. This gives a total of:

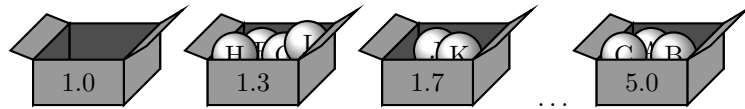
$$\sum_{i=0}^k s_i^{II}(n).$$

1.5.3 $L \rightarrow L$: n Labeled Balls in k Labeled Boxes

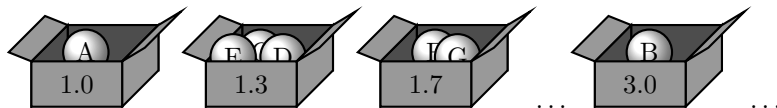
Example 1.12. At the end of the semester, each of the $n = 70$ students of combinatorics will be assigned one of the $k = 11$ grades from $\{1.0, 1.3, \dots, 5.0\}$. Both Torsten and the students think that the outcome



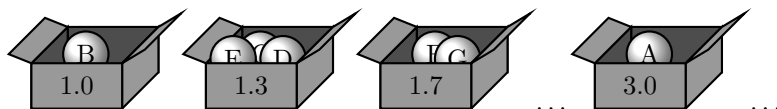
would be preferable to the outcome



so the boxes (grades) are clearly labeled (we could not draw the boxes 2.0 till 4.0 due to lack of space) . Furthermore, Alice (A) and Bob (B) insist that they would notice the difference between the arrangement



and the arrangement



so the balls (students) are labeled as well. Such arrangements correspond directly to functions $f : [n] \rightarrow [k]$, mapping each of the n balls to one of the k boxes, or here: Mapping every student to their grade.

As before, we consider three subcases:

≤ 1 ball per box Of course, this is only possible if there are at most as many balls as boxes ($n \leq k$).

Such arrangements correspond to injective functions $f : [n] \rightarrow [k]$. We first choose the image of $1 \in [n]$ (there are k possibilities), then the image

of $2 \in [n]$ (there are $k - 1$ possibilities left) and so on. Therefore, the number of injective functions (and therefore arrangements with at most one ball per box) is:

$$k \cdot (k - 1) \cdot \dots \cdot (k - n + 1) = \frac{k!}{(k - n)!} = \binom{k}{n} n!.$$

≥ 1 ball per box Of course, this is only possible if there are at least as many balls as boxes ($n \geq k$).

These arrangements correspond to surjective functions from $[n]$ to $[k]$. They can also be thought of as partitions of $[n]$ into k non-empty distinguishable (!) parts. So we count the number of ways to partition $[n]$ into k non-empty indistinguishable parts (there are $s_k^{II}(n)$) and multiply this by the number of ways $k!$ to assign labels to the parts afterwards. So in total, there are

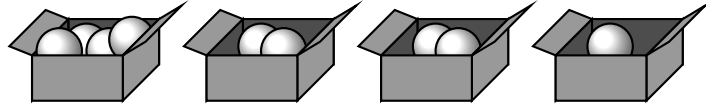
$$k! s_k^{II}(n)$$

ways to assign n labeled balls to k non-empty labeled boxes.



arbitrary number of balls per box There are k choices for each of the n balls, so k^n arrangements in total.

1.5.4 $U \rightarrow U$: n Unlabeled Balls in k Unlabeled Boxes

Example 1.13. Adria used to play tennis but now wants to sell her old tennis balls (n in total) on the local flea market. She wants to sell them in batches since selling them individually is not worth the trouble. So she packs them into k bags. Some people may want to buy bigger batches than others so she figures it might be good to have batches of varying sizes ready. She wonders how many ways there are to pack the balls into bags. One such arrangement ($n = 9, k = 4$) could be:



Balls and boxes are unlabeled. Adria cannot distinguish any two balls and can also not distinguish boxes with the same number of balls.

Even though boxes have no intrinsic ordering, we need to somehow arrange them on this two-dimensional paper. In order to not accidentally think that  and  are different, we use the convention of drawing boxes in decreasing order of balls. With this convention, an arrangement will look different on paper if and only if it is actually different in our sense.

With this in mind we see the number of arrangements of n unlabeled balls in k unlabeled boxes is equal to the number of ways to *partition* the integer n into k non-negative summands. For example:

$$9 = 4 + 2 + 2 + 1.$$

Partitions where merely the order of the summands differ are considered the same, so again we use the convention of writing summands in decreasing order.

≤ 1 ball per box Of course, this is only possible if there are at most as many balls as boxes ($n \leq k$).

In that case, there is only one way to do it: Put every ball in its own box. Then there are n boxes with a ball and $k - n$ empty boxes.

≥ 1 ball per box Of course, this is only possible if there are at least as many balls as boxes ($n \geq k$).

As discussed before, we count the number of ways in which the integer n can be partitioned into exactly k positive parts, i.e.

$$n = a_1 + a_2 + \dots + a_k, \text{ where } a_1 \geq a_2 \geq \dots \geq a_k \geq 1.$$

This is counted by the *partition function*, denoted by $p_k(n)$. A few values are easy to determine:

- $p_0(n) = 0$ (for $n \geq 1$): No positive number can be partitioned into zero numbers.
- $p_n(n) = 1$: To write n as the sum of n positive numbers, there is exactly one choice:

$$n = \underbrace{1 + 1 + \dots + 1}_{n \text{ times}}.$$

To compute other values, observe the following recursion:

$$p_k(n) = p_k(n - k) + p_{k-1}(n - 1).$$

To see this, observe the two cases:

- Either $a_k = 1$. Then $a_1 + \dots + a_{k-1}$ is a partition of $n - 1$ and there are $p_{k-1}(n - 1)$ ways for this.
- Or $a_k \geq 2$. This means each a_i is at least 2 (for $0 \leq i \leq k$). There is a bijection between:
 - (i) Partitions of n into k parts of size at least 2.
 - (ii) Partitions of $n - k$ into k parts of size at least 1.

To go from (i) to (ii), just remove 1 from each part and to go back, just add 1 to each part. We already know how to count (ii), there are $p_k(n - k)$ such partitions.

arbitrary number of balls per box Similar to what we did in the $\mathbf{L} \rightarrow \mathbf{U}$ case, first decide how many boxes i should be non-empty and then count how many arrangements with exactly i non-empty boxes exist:

$$\sum_{i=1}^k p_i(n).$$

This can be thought of as the number of integer partitions of n into *at most* k non-zero parts.

n balls	k boxes	≤ 1 per box	≥ 1 per box	arbitrary
U	L	$\binom{k}{n}$	$\binom{n-1}{k-1}$	$\binom{n+k-1}{k-1}$
L	U	1	$s_k^{II}(n)$	$\sum_{i=1}^k s_i^{II}(n)$
L	L	$\binom{k}{n} n!$	$s_k^{II}(n) k!$	k^n
U	U	1	$p_k(n)$	$q_k(n) = \sum_{i=1}^k p_i(n)$

Table 1: The twelvefold way.

1.5.5 Summary: The Twelvefold Way

The twelve variations of counting arrangements of balls in boxes are called the *Twelvefold Way*. A summary of our results is given in Table 1.

We also summarize the interpretations of n **L**abeled or **U**nabeled balls in k **L**abeled or **U**nabeled boxes:

Arrangements	Correspond to
U \rightarrow L	non-negative Integer solutions of $x_1 + \dots + x_k = n$.
L \rightarrow U	Partitions of the set $[n]$ into k parts.
L \rightarrow L	Functions from $[n]$ to $[k]$
U \rightarrow U	Partitions of the number n into k non-negative integers.

1.6 Binomial Coefficients – Examples and Identities

Example 1.14 (Lattice Paths). A monotone lattice paths in the grid $L := \{0, 1, \dots, m\} \times \{0, 1, \dots, n\}$ is a path starting in the bottom left corner $(0, 0)$ and ending in the top right corner (m, n) , taking single steps upwards or rightwards see Figure 12.

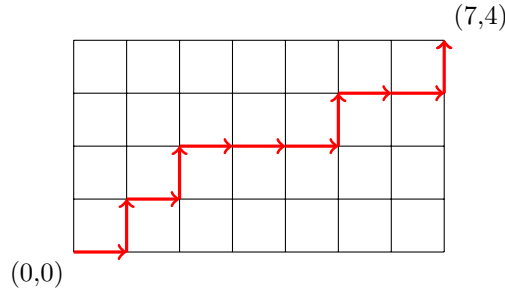


Figure 12: Here $m = 7$ and $n = 4$. The lattice path $\rightarrow \uparrow \rightarrow \uparrow \rightarrow \rightarrow \rightarrow \uparrow \rightarrow \rightarrow \uparrow$ is shown.

In other words, a lattice path is a sequence of “ \uparrow ” (upward steps) and “ \rightarrow ” (rightward steps) with a total of n times \uparrow and m times \rightarrow . In yet other words, the lattice paths correspond to permutations of the multiset $\{m \cdot \rightarrow, n \cdot \uparrow\}$ and by Theorem 1.8 their number is

$$\binom{m+n}{m}.$$

Example (Cake Number). For positive integers m and n the *cake number* $c(m, n)$ is the maximum number of pieces obtained when cutting an m -dimensional cake by n cuts, i.e. the maximum number of connected components of $\mathbb{R}^m \setminus \bigcup_{i=1}^n H_i$, where H_i is an $(m-1)$ -dimensional affine hyperplane. For $m=2$, this simply means: Put n lines into the plane and observe into how many pieces the plane is cut. See Figure 13 for an example.

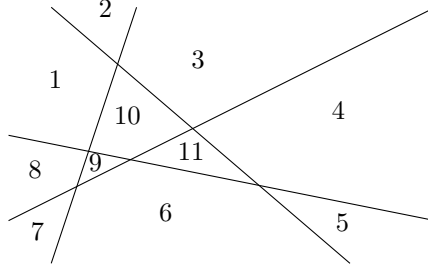


Figure 13: The plane ($m=2$) is cut by $n=4$ lines. In this particular configuration, this gives 11 pieces. No configuration with this n and m can achieve a larger number of pieces.

It turns out that

$$c(m, n) = \sum_{i=0}^m \binom{n}{i}.$$

Here $\binom{n}{i}$ is defined to be zero for $i > n$. For instance

$$c(2, 4) = \binom{4}{0} + \binom{4}{1} + \binom{4}{2} = 1 + 4 + 6 = 11,$$

so the example given in Figure 13 is tight.

The cake numbers arise in a seemingly unrelated setting: For $m \geq 3$, consider any finite set X of n points in \mathbb{R}^{m-1} in general position (no m points on a hyperplane, no $m+1$ on a $(m-2)$ -dimensional sphere). Then the number of subsets Y of X that can be separated from $X \setminus Y$ by a $(m-2)$ -dimensional sphere is equal to $c(m, n)$.

For $m=3$ this means: Consider n points in the plane where no 3 points are on a line and no 4 points on a circle². Then the number of subsets Y of X that can be captured by a circle is equal to $c(3, n)$. Consider Figure 14 for an example. We count all sets that can be captured there.

size 0: The empty set.

size 1: All 5 single element sets.

size 2: All except for $\{b, e\}$ and $\{a, c\}$, so $\binom{5}{2} - 2 = 8$.

size 3: $\{a, b, c\}$, $\{a, b, d\}$, $\{a, b, e\}$, $\{a, d, e\}$, $\{b, c, d\}$, $\{b, d, e\}$, $\{c, d, e\}$, so 7 of them.

size 4: All except for $\{a, b, c, e\}$, so 4 of them.

²These assumptions are very weak: If you pick random points, say by throwing a bunch of darts onto the plane, the probability that the points will be in general position is 1.

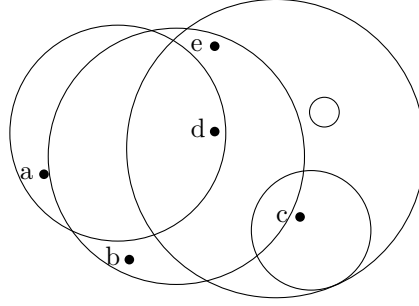


Figure 14: There are $n = 5$ points in the plane ($m = 3$). Some subsets can be captured by a circle (the pictures shows: $\{a, d\}$, $\{b, d, e\}$, $\{c, d, e\}$, $\{c\}$, \emptyset), some cannot, for example $\{b, c, e\}$ (such a circle would also contain d).

size 5: The set $\{a, b, c, d, e\}$.

That is $1 + 5 + 8 + 7 + 4 + 1 = 26$ sets in total, exactly the cake number $c(3, 5) = \binom{5}{0} + \binom{5}{1} + \binom{5}{2} + \binom{5}{3} = 1 + 5 + 10 + 10 = 26$.

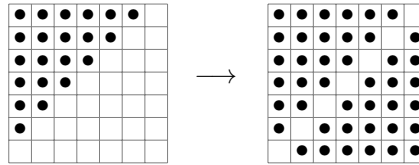
There is a seemingly infinite number of useful identities involving binomial coefficients, we will show a few of them.

Theorem 1.15 (Proofs by Picture). *We have:*

$$(i) \sum_{k=1}^n k = \frac{(n+1)^2 - (n+1)}{2} = \binom{n+1}{2}.$$

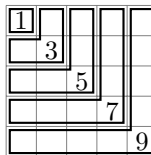
$$(ii) \sum_{k=1}^n (2k - 1) = n^2.$$

Proof. (i) For the first identity, arrange $1 + 2 + \dots + n$ dots in a triangle, mirror it, and observe that this fills all positions of a square of size $(n+1) \times (n+1)$ except for the diagonal. Here is a picture for $n = 6$:



Therefore $2 \cdot \sum_{k=1}^n k = (n+1)^2 - (n+1) = n(n+1)$. Dividing by 2 proves the claim.

(ii) Note how tiles of with sizes of the first n odd numbers can be arranged to form a square of size $n \times n$, here a picture for $n = 5$:



□

Theorem 1.16 (Proofs by Double Counting). *We have*

$$(i) \sum_{k=0}^n 2^k \binom{n}{k} = 3^n.$$

$$(ii) \sum_{i=0}^m \binom{n+i}{i} = \binom{n+m+1}{m}.$$

Proof. (i) We double count strings over the alphabet $\{0, 1, 2\}$ of length n , in other words, the n permutations of the multiset $\{\infty \cdot 0, \infty \cdot 1, \infty \cdot 2\}$ where there is infinite supply of the types 0, 1, 2.

First way: There are three possibilities per character, so 3^n possibilities in total.

Second way: First choose the number of times k that the letter “0” should be used ($0 \leq k \leq n$). Then choose the positions for those characters, there are $\binom{n}{k}$ possibilities. Finally, choose for each of the remaining $(n - k)$ positions if they should be 1 or 2, there are 2^{n-k} choices. So in total, there is this number of possibilities:

$$\sum_{k=0}^n 2^{n-k} \cdot \binom{n}{k}$$

By reversing the order of summation, this is equal to

$$\sum_{k=0}^n 2^k \cdot \binom{n}{n-k} = \sum_{k=0}^n 2^k \cdot \binom{n}{k}.$$

This proves the claim.

(ii) We double count the lattice paths in the lattice $\{0, \dots, m\} \times \{0, \dots, n\}$ of width m and height n .

First way: We already counted them in Example 1.14. There are $\binom{n+m}{m}$ of them.

Second way: Any path must reach the last row (row n) eventually, using exactly one of the “ \uparrow ”-steps from $(k, n-1)$ to (k, n) where $0 \leq k \leq m$. See Figure 15 for a sketch.

There is only one way a lattice path can continue from (k, n) , namely go rightwards until (m, n) . There are $\binom{k+n-1}{k}$ ways to get from $(0, 0)$ to $(k, n-1)$ though, so the number of lattice paths is:

$$\sum_{k=0}^m \binom{k+n-1}{k}.$$

This gives the equality $\binom{n+m}{m} = \sum_{k=0}^m \binom{k+n-1}{k}$. In the claim, we merely replaced n by $n+1$. □

A few other identities can be derived by using the connection of multinomial coefficients to polynomials.

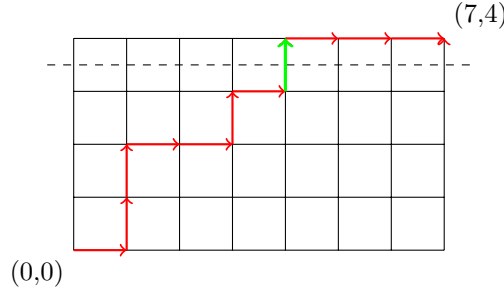


Figure 15: At some point, a lattice path must cross the dashed line, i.e. use one of the edges going to the last row. We count the number of paths using the highlighted edge from $(4, 3)$ to $(4, 4)$: There are $\binom{4+3}{4}$ ways to get from $(0, 0)$ to $(4, 3)$ and one way to get from $(4, 4)$ to $(7, 4)$, so $\binom{4+3}{4} \cdot 1$ paths in total.

Theorem 1.17 (Proofs by Analysis). *We show three equations (i), (ii) and (iii), see below.*

Proof. Start with the binomial formula:

$$(x + y)^n = \sum_{i=0}^n \binom{n}{i} x^i y^{n-i}.$$

Setting $y = 1$ yields:

$$(i) \quad (x + 1)^n = \sum_{i=0}^n \binom{n}{i} x^i.$$

Deriving by x gives:

$$n(x + 1)^{n-1} = \sum_{i=1}^n i \binom{n}{i} x^{i-1}.$$

Now set $x = 1$ and get:

$$(ii) \quad n \cdot 2^{n-1} = \sum_{i=0}^n i \binom{n}{i}.$$

Taking the second derivative of (i) yields:

$$n(n-1)(x+1)^{n-2} = \sum_{i=2}^n i(i-1) \binom{n}{i} x^{i-2}.$$

Again, we set $x = 1$

$$n(n-1)2^{n-2} = \sum_{i=0}^n i(i-1) \binom{n}{i}.$$

Adding (ii) to this gives:

$$(iii) \quad n(n+1)2^{n-2} = \sum_{i=0}^n i^2 \binom{n}{i}.$$

□

1.7 Permutations of Sets

Remember that n -permutations of $[n]$ are bijections $\pi : [n] \rightarrow [n]$. From now on we will just say “permutation of $[n]$ ” and drop the “ n -”.

For such a permutation, the pair (i, j) of two numbers from $[n]$ is called an *inversion* if the numbers are ordered, but swapped by the permutation, i.e.

$$(i, j) \text{ is inversion} \Leftrightarrow i < j \wedge \pi^{-1}(i) > \pi^{-1}(j).$$

Take for example the permutation $\pi = 31524$. It has the inversions $(1, 3)$, $(2, 3)$, $(2, 5)$, $(4, 5)$.

For $i \in [n]$ define $\alpha_i := |\{j \in [n] \mid (i, j) \text{ is inversion}\}|$. The *inversion sequence* of π is the sequence $\alpha_1 \alpha_2 \dots \alpha_n$.

The *inversion number* or *disorder* of a permutation π is the number of its inversions, so $\alpha_1 + \cdots + \alpha_n$. For our example the inversion sequence is $1, 2, 0, 1, 0$ and its disorder is 4.

Since for any $i \in [n]$, there are only $n-i$ numbers bigger than i , any inversion sequence $\alpha_1, \dots, \alpha_n$ satisfies:

$$0 \leq \alpha_i \leq n - i \quad (i \in [n]). \quad (\star)$$

There are exactly $n!$ integer sequences satisfying (\star) (n choices for α_1 , $n-1$ choices for $\alpha_2 \dots$), just as many as there are permutations. This is no coincidence: We now show that each permutation has a distinct such sequence as an inversion sequence. By cardinality we then also know that each such sequence is the inversion sequence of some permutation, i.e. there is a one-to-one correspondence.

Theorem 1.18. *Any permutation can be recovered from its inversion sequence.*

Proof. Let $\alpha = (\alpha_1, \dots, \alpha_n)$ be integers satisfying (\star) . We will construct a permutation π with this inversion sequence and it will be clear that we will not have any freedom in the construction of π , i.e. π will be the *unique* permutation with inversion sequence α .

In the beginning, π has n free positions still to be determined:

$$\pi = \underbrace{\hspace{10em}}_{n \text{ positions}}$$

We determine the position containing 1, i.e. we determine $\pi^{-1}(1)$: Remember that α_1 is the number of elements that are bigger than 1 but arranged to the left of 1 by π . Since *every element* is bigger than 1, α_1 is just the number of elements arranged to the left of 1, in other words $\pi^{-1}(1) = \alpha_1 + 1$.

Take for instance the inversion sequence $\alpha = (5, 3, 4, 0, 2, 1, 0, 0)$. Then from $\alpha_1 = 5$ we know that 1 is in the sixth position like this:

$$\pi = \frac{1}{\dots}$$

Now recall that α_2 is the number of elements bigger than 2 that are sorted to the left of 2, this means, α_2 is exactly the number of (currently) unoccupied positions to the left of 2. In the example we have $\alpha_2 = 3$, so π must look like this:

$$\pi = \frac{2}{1}$$

For general: $i \in [n]$, if we have already placed all number from 1 to $i - 1$, we can derive that $\pi^{-1}(i)$ must be the position of the free slot that has exactly α_i free slots to the left of it.

In the example, since $\alpha_3 = 4$ we put 3 to the unoccupied position that has four unoccupied positions to the left of it:

$$\pi = _ _ _ \underline{2} _ \underline{1} \underline{3} _$$

Continuing this way, we can reconstruct π completely:

$$\pi = \underline{4} \underline{7} \underline{6} \underline{2} \underline{5} \underline{1} \underline{3} \underline{8} . \quad \square$$

1.7.1 Generating Permutations

To generate all permutations of $[n]$ use the following algorithm.

Start: Write $1, 2, \dots, n$ with left arrows over them.

- Step:
- Locate the largest mobile integer m , i.e., an integer such that it arrows to a smaller neighbouring number.
 - Swap m with that neighbour.
 - Change the direction of the arrows on all integers p , $p > m$.

Let us illustrate this method with the following example.

$$\begin{array}{cccc} \overleftarrow{1} & \overleftarrow{2} & \overleftarrow{3} & \overleftarrow{4} \\ \overleftarrow{1} & \overleftarrow{2} & \overleftarrow{4} & \overleftarrow{3} \\ \overleftarrow{1} & \overleftarrow{4} & \overleftarrow{2} & \overleftarrow{3} \\ \overleftarrow{4} & \overleftarrow{1} & \overleftarrow{2} & \overleftarrow{3} \end{array} \quad \begin{array}{ccc} \overleftarrow{1} & \overleftarrow{2} & \overleftarrow{3} \\ & & \end{array}$$

$$\begin{array}{cccc} \overrightarrow{4} & \overleftarrow{1} & \overleftarrow{3} & \overleftarrow{2} \\ \overleftarrow{1} & \overrightarrow{4} & \overleftarrow{3} & \overleftarrow{2} \\ \overleftarrow{1} & \overleftarrow{3} & \overrightarrow{4} & \overleftarrow{2} \\ \overleftarrow{1} & \overleftarrow{3} & \overleftarrow{2} & \overrightarrow{4} \end{array} \quad \begin{array}{ccc} \overleftarrow{1} & \overleftarrow{3} & \overleftarrow{2} \\ & & \end{array}$$

$$\begin{array}{cccc} \overleftarrow{3} & \overleftarrow{1} & \overleftarrow{2} & \overleftarrow{4} \\ \overleftarrow{3} & \overleftarrow{1} & \overleftarrow{4} & \overleftarrow{2} \\ \overleftarrow{3} & \overleftarrow{4} & \overleftarrow{1} & \overleftarrow{2} \\ \overleftarrow{4} & \overleftarrow{3} & \overleftarrow{1} & \overleftarrow{2} \end{array} \quad \begin{array}{ccc} \overleftarrow{3} & \overleftarrow{1} & \overleftarrow{2} \\ & & \end{array}$$

$$\begin{array}{cccc} \overrightarrow{4} & \overrightarrow{3} & \overleftarrow{2} & \overleftarrow{1} \\ \overrightarrow{3} & \overrightarrow{4} & \overleftarrow{2} & \overleftarrow{1} \\ \overrightarrow{3} & \overleftarrow{2} & \overrightarrow{4} & \overleftarrow{1} \\ \overrightarrow{3} & \overleftarrow{2} & \overleftarrow{1} & \overrightarrow{4} \end{array} \quad \begin{array}{ccc} \overrightarrow{3} & \overleftarrow{2} & \overleftarrow{1} \\ & & \end{array}$$

$$\begin{array}{cccc} \overleftarrow{2} & \overrightarrow{3} & \overleftarrow{1} & \overleftarrow{4} \\ \overleftarrow{2} & \overrightarrow{3} & \overleftarrow{4} & \overleftarrow{1} \end{array} \quad \begin{array}{ccc} \overleftarrow{2} & \overrightarrow{3} & \overleftarrow{1} \\ & & \end{array}$$

$$\begin{array}{cccc} \overleftarrow{2} & \overleftarrow{4} & \overrightarrow{3} & \overleftarrow{1} \\ \overleftarrow{4} & \overleftarrow{2} & \overrightarrow{3} & \overleftarrow{1} \end{array}$$

$$\begin{array}{cccc} \overrightarrow{4} & \overleftarrow{2} & \overleftarrow{1} & \overrightarrow{3} \\ \overleftarrow{2} & \overrightarrow{4} & \overleftarrow{1} & \overrightarrow{3} \\ \overleftarrow{2} & \overleftarrow{1} & \overrightarrow{4} & \overrightarrow{3} \\ \overleftarrow{2} & \overleftarrow{1} & \overrightarrow{3} & \overrightarrow{4} \end{array} \quad \overleftarrow{2} \quad \overleftarrow{1} \quad \overrightarrow{3}$$

1.7.2 Cycle Decompositions

Remark. The lecture mentioned a puzzle presented on the Youtube Channel MinutePhysics. It *may* have something to do with cycles (but we do not want to spoil the puzzle for you), feel free to check it out on

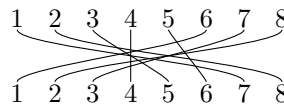
<http://youtube.com/watch?v=eivGIBKIK6M>

We already know several ways to specify a permutation.

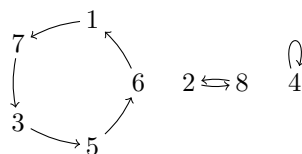
- Write it as a string: $\pi = 78546132$.
- Write it explicitly as a map:

$$\begin{array}{c|cccccccc} i & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \hline \pi(i) & 7 & 8 & 5 & 4 & 6 & 1 & 3 & 2 \end{array}$$

- Another way to specify the map would be to connect i with $\pi(i)$, maybe like this:



- In the last drawing every number occurs twice (once as a preimage and once as an image), now we only take one copy of every number and give a direction to the edges (from preimage to image):



In other words, we draw an edge from $i \rightarrow j$ if $\pi(i) = j$. Since π is a map, every node gets one outgoing edge and since π is bijective, every node gets exactly one incoming edge. From this it is easy to see that the image we get is a collection of cycles: When starting at any node i and following the arrows, i.e. walking along the path $i \rightarrow \pi(i) \rightarrow \pi(\pi(i)) \rightarrow \pi(\pi(\pi(i))) \rightarrow \dots$, we must at some point – since there are only finitely many elements – come for the first time to an element where we already were. This must be i , since all other nodes already have an incoming edge. So i was indeed in a cycle.

Instead of drawing the picture we prefer to just list the cycles like this:

$$\pi = (17356)(28)(4)$$

where cycles are group by “()” and elements within a cycle are given in order. The representation at hand is called a *disjoint cycle decomposition*. Note that every element occurs in exactly one cycle (some cycles may have length 1). Note that this composition is, strictly speaking, not unique, we could also write $\pi = (4)(73561)(28)$, where we changed the order of two cycles and “rotated” 7 to the beginning of its cycle. In the following, we will not distinguish disjoint cycle decompositions that differ only in these ways.

Before we generalize disjoint cycle decompositions to arbitrary cycle decompositions, we define the *product* of two permutations:

If $\pi_1, \pi_2 \in S_n$ are permutations, the product $\pi_1 \cdot \pi_2$ (or just $\pi_1\pi_2$ for short) is the permutation π with $\pi(i) = \pi_2(\pi_1(i))$ (i.e. first apply π_1 , then π_2). Note, that when viewed as maps: $\pi_1 \cdot \pi_2 = \pi_2 \circ \pi_1$.

Now if we write a single cycle, e.g. $\pi \in S_7$, $\pi = (137)$, we mean the permutation that permutes the elements in the cycle along the cycle (here $\pi(1) = 3$, $\pi(3) = 7$, $\pi(7) = 1$) and leaves all other elements in place, in our example π is the permutation with the disjoint cycle decomposition $(137)(2)(4)(5)(6)$.

We can now talk about cycle decompositions where elements may occur more than once, e.g. $\pi = (136)(435)(7452)(6)(23)$. It is the product of the involved cycles $((136), (435), \dots)$.

If $\pi \in S_n$ is a permutation given as a cycle decomposition, then π can be evaluated by *parsing*, where parsing an element $i \in [n]$ in a cycle decomposition means

- Have a current element c , initially $c = i$.
- Go through cycles from left to right.
 - If the current element c is part of a cycle, replace c with the next element in this cycle.
- In the end of this process $\pi(i)$ is the current element c .

Take for instance $\pi = (136)(435)(7452)(6)(23)$ and $i = 3$. Then we start with the current element $c = 3$. The first cycle (136) contains $c = 3$, so we change c to 6 and go on. The next cycle (435) does not contain $c = 6$, and neither does (7452) so we move past these cycles without changing c . The next cycle (6) contains $c = 6$ but does not alter it. We therefore end with $\pi(3) = 6$. As another example, consider how $\pi(1)$ is evaluated:

$$\begin{array}{cccccc} c=1 & c=3 & c=5 & c=2 & c=2 & c=3 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ (136) & (435) & (7452) & (6) & (23) & \end{array}$$

So $\pi(1) = 3$.

Theorem 1.19. *If $\pi_1, \pi_2 \in S_n$ are permutations given in cycle decomposition, then a cycle decomposition of $\pi_1 \cdot \pi_2$ is given by concatenating the cycle decompositions of π_1 and π_2 .*

Proof. Let the cycle decompositions be $\pi_1 = C_1 C_2 \dots C_k$, $\pi_2 = C'_1 C'_2 \dots C'_l$. When parsing $C_1 C_2 \dots C_k C'_1 C'_2 \dots C'_l$ with $i \in [n]$ then the current element after $C_1 C_2 \dots C_k$ will be $\pi_1(i)$ and therefore, after going through the remaining cycles $C'_1 \dots C'_l$, the current element will be $\pi_2(\pi_1(i)) = (\pi_1 \cdot \pi_2)(i)$. So the concatenation is indeed a cycle decomposition of $\pi_1 \cdot \pi_2$. \square

Note that non-disjoint cycle decompositions do not always commute, e.g.

$$(435)(321) = (13542) \neq (15432) = (321)(435)$$

where the identities can be verified by parsing. The claims of the following theorem are easy and we will skip the proofs:

Theorem 1.20. *Let $\pi \in S_n$ be given in cycle decomposition.*

- *Swapping adjacent cycles has no effect if they are disjoint (i.e. no number occurs in both cycles), e.g. $(13)(524) = (524)(13)$.*
- *A cycle decomposition of π^{-1} is obtained by swapping the order of all cycles and reversing the elements in each cycle, e.g. $((321)(435))^{-1} = (534)(123)$.*
- *Cyclic shifts in a cycle have no effect. $(534) = (345) = (453)$.*
- *Up to cyclic shifts and order of cycles, there is a unique decomposition into disjoint cycles.*

Let $\text{id} \in S_n$ be the identity permutation ($\text{id}(i) = i$ for $i \in [n]$). Define the *order* of $\pi \in S_n$ to be the smallest k such that

$$\pi^k = \underbrace{\pi \cdot \pi \cdot \dots \cdot \pi}_{k \text{ times}} = \text{id}.$$

Theorem 1.21. *The order of π is the least common multiple of the lengths of the cycles in the disjoint cycle decomposition of π .*

Proof. Assume π is composed of the disjoint cycles $C_1 C_2 \dots C_m$ and an element $i \in [n]$ is contained in a cycle C_j of length l . Then for any $k \in \mathbb{N}$, the value $\pi^k(i)$ is obtained by parsing i through i

$$\underbrace{C_1 C_2 \dots C_m C_1 C_2 \dots C_m \dots C_1 C_2 \dots C_m}_{k \text{ copies of } C_1 \dots C_m}$$

Only the copies of C_j can affect the current value if we start with i , so the result is the same as when parsing i through k copies of just C_j . From this we see that $\pi^k(i) = i$ if and only if k is a multiple of l . This shows that $\pi^k = \text{id}$ if and only if k is a common multiple of the cycle lengths. So, the order of π is the least common multiple of them. \square

1.7.3 Transpositions

A cycle of length 2 is a *transposition*.

Define *discriminant* of π as $N(\pi) := n - \#C$ where $\#C$ is the number of cycles in a disjoint cycle decomposition of π . Note that we may not omit single element cycles now, they count towards $\#C$:

In S_5 we have for example $N(\text{id}) = N((1)(2)(3)(4)(5)) = n - 5 = 0$ and $N((134)(25)) = 5 - 2 = 3$.

Theorem 1.22.

(i) Any permutation can be written as the product of transpositions.

(ii) If π is the product of k transpositions ($k \in \mathbb{N}$), then $N(\pi)$ and k have the same parity (i.e. $N(\pi) \equiv k \pmod{2}$).

Proof. (i) Let π be any permutation. We already know that we can write π as a product of cycles. So to show that π can be written as a product of transpositions, it suffices to show that any cycle can be written as the product of transpositions.

Let $C = (a_1 a_2 \dots a_l)$ be a cycle of length l (assume $l \geq 2$, since cycles of length 1 correspond to identity permutations and can be omitted from any cycle decomposition).

Then we can write C as the product of $l - 1$ transpositions:

$$C = (a_1 a_2 \dots a_l) = (a_{l-1} a_l)(a_{l-2} a_{l-1}) \dots (a_2 a_3)(a_1 a_2)$$

Verify this by parsing: If $i \neq l$, then parsing a_i yields:

$$\begin{array}{ccccccccccc} c=a_i & c=a_i & & c=a_i & c=a_i & c=a_{i+1} & c=a_{i+1} & c=a_{i+1} & c=a_{i+1} & c=a_{i+1} & c=a_{i+1} \\ \downarrow & \downarrow & & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ (a_{l-1} a_l) & (a_{l-2} a_{l-1}) & \dots & (a_i a_{i+1}) & (a_{i-1} a_i) & \dots & (a_2 a_3) & (a_1 a_2) & & & \end{array}$$

So parsing sends a_i to a_{i+1} as desired. When parsing a_l we have:

$$\begin{array}{ccccccc} c=a_l & c=a_{l-1} & c=a_{l-2} & c=a_3 & c=a_2 & c=a_1 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ (a_{l-1} a_l) & (a_{l-2} a_{l-1}) & \dots & (a_2 a_3) & (a_1 a_2) & \end{array}$$

so we get a_1 as desired.

(ii) Let π be any permutation with k cycles in a disjoint cycle decomposition. Let (ab) be any transposition. We determine the number of cycles in a disjoint cycle decomposition of $(ab) \cdot \pi$ depending on a and b .

Case 1: a and b are part of different cycles in π . Then there is a disjoint cycle decomposition of π looking like this:

$$\pi = (aX)(bY)C_3 \dots C_k$$

where X and Y are sequences of elements and C_3, \dots, C_k are cycles. By parsing we verify:

$$(ab) \cdot \pi = (ab)(aX)(bY)C_3 \dots C_k = (aYbX)C_3 \dots C_k.$$

The last term is a disjoint cycle decomposition of $(ab) \cdot \pi$ into $k - 1$ cycles.

Case 2: a and b are part of the same cycle in π . Then there is a disjoint cycle decomposition of π looking like this:

$$\pi = (aXbY)C_2 \dots C_k$$

where X and Y are sequences of elements and C_2, \dots, C_k are cycles. By parsing we verify:

$$(ab) \cdot \pi = (ab)(aXbY)C_2 \dots C_k = (aY)(bX)C_2 \dots C_k.$$

The last term is a disjoint cycle decomposition of $(ab) \cdot \pi$ into $k + 1$ cycles.

In both cases, adding a transposition changed the number of cycles by one, which means that the parity of the number of cycles changed. Therefore, the claim follows by induction. \square

We call the number $s_k^I(n)$ of n -permutations of $[n]$ with exactly k cycles in a disjoint cycle decomposition the *unsigned Stirling number of first kind*. More precisely $s_k^I(n) = |\{\pi \in S_n \mid N(\pi) = n - k\}|$.

Theorem 1.23. *For all $n, k \geq 1$*

- (i) $s_0^I(0) = 1$ and $s_k^I(0) = s_0^I(n) = 0$,
- (ii) $s_k^I(n) = (n - 1)s_k^I(n - 1) + s_{k-1}^I(n - 1)$.

Proof. (i) The empty permutation is the only permutation of 0 elements. Thus $s_0^I(0) = 1$ and $s_k^I(0) = 0$, since the empty permutation has no cycle and $k \geq 1$.

On the other hand every permutation of $n \geq 1$ elements has at least one cycle and hence $s_0^I(n) = 0$.

- (ii) Let $S_n^k \subset S_n$ denote the set of permutations in S_n with exactly k cycles in a disjoint cycle decomposition. Then $|S_n^k| = s_k^I(n)$. We define a map $\phi : S_n^k \rightarrow S_{n-1}^k \cup S_{n-1}^{k-1}$ as follows. Consider a permutation $\pi \in S_n^k$ and let (AnB) denote the cycle containing n (A and B may be empty). If $(AnB) = (n)$, remove the entire cycle (n) . Otherwise replace the cycle (AnB) with (AB) . In the first case there are $k - 1$ cycles in a disjoint cycle decomposition of $\phi(\pi)$. In the latter case $\phi(\pi)$ still has k cycles in a disjoint cycle decomposition.

Consider a disjoint cycle decomposition of some $\sigma \in S_{n-1}^k \cup S_{n-1}^{k-1}$. We count how many permutations $\pi \in S_n^k$ are mapped to σ .

Case 1: $\sigma \in S_{n-1}^{k-1}$. Then the only permutation in S_n^k mapped to σ is obtained by adding the cycle (n) to σ .

Case 2: $\sigma \in S_{n-1}^k$. In this case we may place n between any element $x \in [n - 1]$ and its successor in σ (i.e., put it right behind x in the cycle containing x). Then n has different positions for distinct choices of $x \in [n - 1]$. Hence $n - 1$ distinct elements from S_n^k are mapped to σ .

Combining these two cases we see that $|S_n^k| = (n - 1)|S_{n-1}^k| + |S_{n-1}^{k-1}|$. \square

Recall that the Stirling numbers of second kind $s_k^{II}(n)$ count the number of partitions of $[n]$ into k non-empty parts. For these numbers we had the formula $s_k^{II}(n) = ks_k^{II}(n - 1) + s_{k-1}^{II}(n - 1)$. There is also a similar picture of the recursion, see Figure 16. From the picture we easily see that $s_1^I(n) = (n - 1)!$ (for $n \geq 1$). This is the number of different cycles of length n formed by n elements.

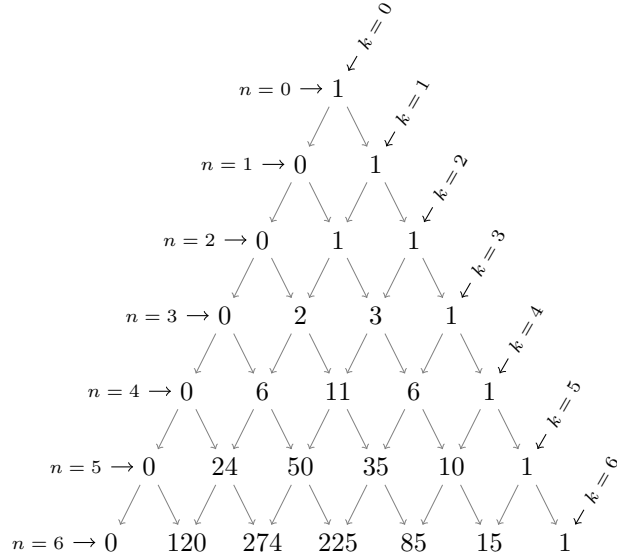


Figure 16: A number $s_k^I(n)$ is obtained as the sum of the number $s_{k-1}^I(n-1)$ toward the top left and $(n-1)$ times the number $s_k^I(n-1)$ towards the top right. E.g. for $n=6, k=3$: $225 = 50 + 5 \cdot 35$.

1.7.4 Derangements

A permutation $\pi \in S_n$ is a *derangement* if it is fixpoint-free, i.e., $\forall i \in [n] : \pi(i) \neq i$. Note that this means that each cycle has length at least 2. The set of all derangements of $[n]$ is denoted by D_n .

We close this chapter by “counting” derangements in Theorem 1.24. This is also meant to demonstrate that there are several different ways to count. Depending on the purpose, different counts may be more or less helpful.

Remark. You might find the notation $|D_n| = !n$ in the literature (sic!). Since we think this notation is too confusing we won’t use it here.

Before we start observe that $|D_1| = 0$ and $|D_2| = 1$. We may also agree on $|D_0| = 1$, since the empty permutation has no fixpoint, but we try to avoid using D_0 .

Theorem 1.24. For a natural number $n \geq 1$ we have

$$|D_n| = (n-1)(|D_{n-1}| + |D_{n-2}|), \text{ if } n \geq 2, \quad (\text{Recursion (i)})$$

$$|D_n| = n|D_{n-1}| + (-1)^n, \quad (\text{Recursion (ii)})$$

$$n! = \sum_{k=0}^n \binom{n}{k} |D_k|, \quad |D_n| = n! - \sum_{k=0}^{n-1} \binom{n}{k} |D_k|, \quad (\text{Recursion (iii)})$$

$$|D_n| = \sum_{k=0}^n \binom{n}{k} (-1)^{n+k} k! = n! \sum_{k=0}^n \frac{(-1)^k}{k!}, \quad (\text{Summation})$$

$$|D_n| = \left\lfloor \frac{n!}{e} + \frac{1}{2} \right\rfloor, \quad (\text{Explicit})$$

$$|D_n| \sim \sqrt{2\pi n} \frac{n^n}{e^{n+1}} \quad (\sim c^{n \log n} \text{ for some } c \in \mathbb{R}), \quad (\text{Asymptotic})$$

$$\sum_{n \geq 0} \frac{|D_n|}{n!} x^n = \frac{e^{-x}}{1-x} \quad \forall x \in \mathbb{R} \setminus \{1\}. \quad (\text{Generating Function})$$

Proof.

Recursion (i): We prove this statement using a map $\phi : D_n \rightarrow D_{n-1} \cup D_{n-2}$ as follows. Consider a derangement $\pi \in D_n$ and let $x, y \in [n]$ be the unique numbers such that $\pi(x) = n$ and $\pi(n) = y$. Written as a string, π is of the form $\pi = AnBy$ where A is a sequence of $x-1$ elements and B is a sequence of $n-x-1$ elements.

If (**case 1**) $x \neq y$, then define $\phi(\pi) := AyB$. Note that $\sigma = \phi(\pi) \in D_{n-1}$, since $\sigma(x) = y \neq x$. For example, if $\pi = 45123$, then $x = 2, y = 3$ and $\phi(\pi) = 4312$.

If (**case 2**) $x = y$, then define $\phi(\pi) := A^*B^*$ where A^* and B^* are identical to A and B except that all numbers bigger than x are decreased by 1. We claim that $\sigma = \phi(\pi) \in D_{n-2}$. If $i < x$, then $\sigma(i) = \phi(i) \neq i$ if $\phi(i) < x$, and $\sigma(i) = \phi(i) - 1 \geq x > i$ if $\phi(i) > x$. Otherwise $i \geq x$ and $\sigma(i) = \phi(i+1) < x < i$ if $\phi(i+1) < x$, and $\sigma(i) = \phi(i+1) - 1 \neq i+1-1 = i$ if $\phi(i+1) > x$. For example, if $\pi = 45132$, then $x = y = 2, AB = 413$ and $\phi(\pi) = 312$.

Claim. Each $\sigma \in D_{n-1} \cup D_{n-2}$ is the image of exactly $(n-1)$ permutations $\pi \in D_n$.

Case 1: $\sigma \in D_{n-1}$. Then there are $n-1$ choices to pick a position $x \in [n-1]$. Let $\sigma(x) = y$, write $\sigma = AyB$ and define $\pi = AnBy$. Then $\phi(\pi) = \sigma$. Note that $\pi \in D_n$, since $\pi(x) = n \neq x$.

Case 2: $\sigma \in D_{n-2}$. Then there are $n-1$ choices to put n on position x (as the first element, into a gap or after the last element). Then there is a unique way to increase $\sigma(i)$ by 1 for each i with $\sigma(i) \geq x$. Finally put x on position n . We obtain $\pi = AnBx$, which is mapped to σ . Note that $\pi \in D_n$ due to similar arguments as above.

Recursion (ii): We will apply induction on n . An induction basis is given by $|D_2| = 1$ and $|D_1| = 0$. Suppose $n \geq 3$. We rewrite Recursion (i) as $|D_n| - n|D_{n-1}| = -(|D_{n-1}| - (n-1)|D_{n-2}|)$. By induction hypothesis, the right side is equal to $-(-1)^{n-1} = (-1)^n$ which proves the claim.

Remark. For the number $|S_n| = P(n) = n!$ of permutations of $[n]$ we have a similar recursion $|S_n| = (n-1)(|S_{n-1}| + |S_{n-2}|)$, since $n! = (n-1)((n-1)! + (n-2)!) = n|S_{n-1}|$.

Recursion (iii): We count all permutations in S_n . On the one hand $|S_n| = n!$. On the other hand each $\pi \in S_n$ induces a derangement on $[n] \setminus F(\pi)$, where $F(\pi)$ is the set of all fixpoints of π . Thus there are $\binom{n}{k}|D_{n-k}|$ different permutations in S_n with exactly k fixpoints each. Hence $n! = |S_n| = \sum_{k=0}^n \binom{n}{k}|D_{n-k}|$.

summation: Proof by induction on n . If $n = 1$, then $|D_1| = 0 = 1 \sum_{k=0}^1 \frac{(-1)^k}{k!} = 1(1-1)$. This gives an induction basis.

Consider $n \geq 2$. Then, using the induction hypothesis (IH):

$$\begin{aligned} |D_n| &= n|D_{n-1}| + (-1)^n \stackrel{\text{IH}}{=} n(n-1)! \sum_{k=0}^{n-1} \frac{(-1)^k}{k!} + (-1)^n \\ &= n! \sum_{k=0}^{n-1} \frac{(-1)^k}{k!} + \frac{n!}{n!} (-1)^n = n! \sum_{k=0}^n \frac{(-1)^k}{k!}. \end{aligned}$$

explicit: Recall that $e^z = \sum_{k \geq 0} \frac{z^k}{k!}$. Then

$$\lim_{n \rightarrow \infty} \frac{|D_n|}{|S_n|} = \lim_{n \rightarrow \infty} \frac{n! \sum_{k=0}^n \frac{(-1)^k}{k!}}{n!} = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{(-1)^k}{k!} = e^{-1} = \frac{1}{e}.$$

So for sufficiently large n we have $|D_n| \sim \frac{n!}{e}$. Applying standard arguments for the speed of convergence of series we obtain $|D_n| = \lfloor \frac{n!}{e} + \frac{1}{2} \rfloor$.

asymptotic: This follows immediately from Stirling's formula $n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$ and the explicit result.

generating function: Let $F(x) := \sum_{n \geq 0} \frac{|D_n|}{n!} x^n$. Then the derivative of F is given by $F'(x) = \sum_{n \geq 1} \frac{|D_n|}{n!} n x^{n-1} = \sum_{n \geq 1} \frac{|D_n|}{(n-1)!} x^{n-1}$. From Recursion (i) we have $\frac{|D_{n+1}|}{n!} - \frac{|D_n|}{(n-1)!} = \frac{|D_{n-1}|}{(n-1)!}$ and thus

$$\begin{aligned} (1-x)F'(x) &= \sum_{n \geq 1} \frac{|D_n|}{(n-1)!} (x^{n-1} - x^n) \\ &= \sum_{n \geq 1} \left(\frac{|D_{n+1}|}{n!} - \frac{|D_n|}{(n-1)!} \right) x^n \\ &= \sum_{n \geq 1} \frac{|D_{n-1}|}{(n-1)!} x^n \\ &= xF(x). \end{aligned}$$

This differential equation with $F(0) = 1$ is solved by $F(x) = \frac{e^{-x}}{1-x}$. \square

2 Inclusion-Exclusion-Principle and Möbius Inversion

In the last theorem of the previous chapter, and in several other places, we have seen summations with alternating signs. This chapter will deal with such kind of results. Consider a finite set X and some properties P_1, \dots, P_m such that for each $x \in X$ and each $i \in [m]$ it is known whether x has property P_i or not. We are interested in the number of elements from X satisfying none of the properties P_i .

Example. Let X be the set of students in the room and let P_1 and P_2 be the properties “being male” and “studies math”, respectively. Suppose there are 36 students, 26 of which are male and 32 of which study math. Among the male students 23 study math. We are interested in the number of students which are neither male nor study math. Let X_1, X_2 denote the set of male students and the set of math students, respectively. Then

$$|X \setminus (X_1 \cup X_2)| = |X| - |X_1| - |X_2| + |X_1 \cap X_2| = 36 - 26 - 32 + 23 = 1.$$

See Figure 17.

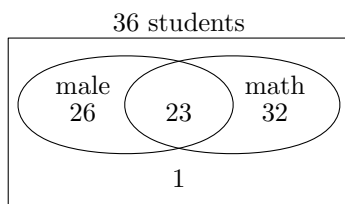


Figure 17: When there are 26 male and 32 math students among 36 students in the class, and 23 of the male students study math, then there is exactly 1 female student who does not study math.

2.1 The Inclusion-Exclusion Principle

Let P_1, \dots, P_m be some m properties and X be an n -element set, where for each $x \in X$ and each $i \in \{1, \dots, m\}$ we have that x has property P_i or x does not have property P_i . For a subset S of properties, let $N(S)$ denote the set of elements in X that have properties P for all $P \in S$. Note that $N(\emptyset) = X$. Clearly, a property P_i just corresponds to the subset X_i of elements of X having this property and $N(S) = \bigcap_{i \in S} X_i$. However, taking properties instead of sets has the advantage that we can take the same subset more than once.

Theorem 2.1 (Inclusion-Exclusion Principle).

Let X be a finite set and P_1, \dots, P_m properties. Further define for $S \subseteq [m]$ the set $N(S) = \{x \in X \mid \forall i \in S : x \text{ has } P_i\}$, i.e. the set of all elements from X having property P_i for all $i \in S$. The number of elements of X that satisfy none of the properties P_1, \dots, P_m is given by

$$\sum_{S \subseteq [m]} (-1)^{|S|} |N(S)|. \quad (1)$$

Proof. Consider any $x \in X$. If $x \in X$ has none of the properties, then $x \in N(\emptyset)$ and $x \notin N(S)$ for any $S \neq \emptyset$. Hence x contributes 1 to the sum (1).

If $x \in X$ has exactly $k \geq 1$ of the properties, call this set of properties $T \in \binom{[m]}{k}$. Then $x \in N(S)$ iff $S \subseteq T$.

The contribution of x to the sum (1) is $\sum_{S \subseteq T} (-1)^{|S|} = \sum_{i=0}^k \binom{k}{i} (-1)^i = 0$.

In the last step we used that for any $y \in \mathbb{R}$ we have $(1-y)^k = \sum_{i=0}^k \binom{k}{i} (-y)^i$ which implies (for $y = 1$) that $0 = \sum_{i=0}^k \binom{k}{i} (-1)^i$. \square

The previous theorem can also be proved inductively.

In some settings we might be interested in a set of elements having a certain set of properties A and none of the properties from a set of properties B . We can handle this by considering the opposites \bar{A} of the properties from A and searching for the elements having none of the properties from \bar{A} and B . In other settings we might be interested in the number of elements satisfying at least one of the properties. The following corollary answers this question.

Corollary 2.2. *The number of elements of X that have at least one of the properties P_1, \dots, P_m is given by*

$$|X| - \sum_{S \subseteq [m]} (-1)^{|S|} N(S) = \sum_{\emptyset \neq S \subseteq [m]} (-1)^{|S|-1} N(S).$$

Given a set X and three properties P_1, P_2, P_3 , the Inclusion-Exclusion-Principle states that the number of elements not in $P_1 \cup P_2 \cup P_3$ is:

$$|X| - |P_1| - |P_2| - |P_3| + |P_1 \cap P_2| + |P_1 \cap P_3| + |P_2 \cap P_3| - |P_1 \cap P_2 \cap P_3|.$$

Figure 18 illustrates this: For example, the number of elements in $P_1 \setminus (P_2 \cup P_3)$

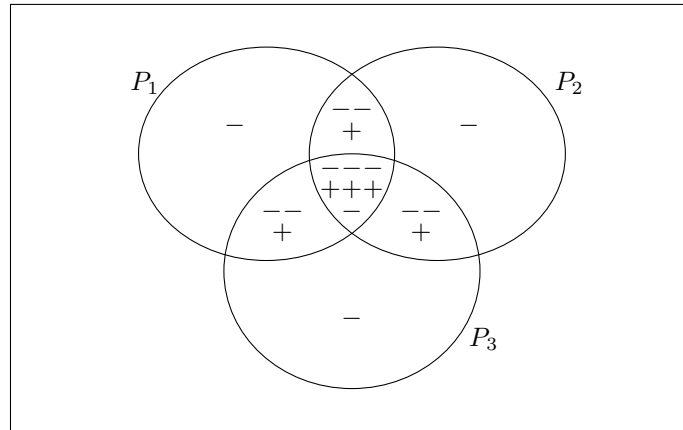


Figure 18: An illustration of the inclusion-exclusion principle for three properties.

is subtracted once from the total number of elements. The number of elements in $(P_1 \cap P_2) \setminus P_3$ is subtracted twice in the beginning (since the elements are in P_1 as well as in P_2) and then added back once. The number of elements

in $P_1 \cap P_2 \cap P_3$ is subtracted three times in the beginning, then added back three times and finally subtracted yet again. The same holds for the other intersections. Altogether one can see that each element in $P_1 \cup P_2 \cup P_3$ contributes exactly -1 to the total sum.

2.1.1 Applications

In the following we apply the principle of inclusion and exclusion (PIE) to count things. The arguments have a fairly rigid pattern:

- (i) **Define “bad” properties:** We identify the things to count as the elements of some universe X except for those having at least one of a set of properties P_1, \dots, P_m . The corresponding sets are denoted by X_1, \dots, X_m (i.e. X_i is the set of elements having property P_i). Given this reformulation of our problem we want to count $X \setminus (X_1 \cup \dots \cup X_m)$.
- (ii) **Count $N(S)$:** For each $S \subseteq [m]$, determine $N(S)$, the number of elements of X having all bad properties P_i for $i \in S$.
- (iii) **Apply PIE:** Apply Theorem 2.1, i.e. the principle of inclusion and exclusion. This yields a closed formula for $|X \setminus (X_1 \cup \dots \cup X_m)|$, typically with one unresolved summation sign.

Theorem 2.3 (Surjections). *The number of surjections from $[k]$ to $[n]$ is:*

$$\sum_{i=0}^n (-1)^i \binom{n}{i} (n-i)^k.$$

Proof. **Define bad properties:** Let X be the set of all maps from $[k]$ to $[n]$. Define the “bad” property P_i for $i \in [n]$ as “ i is not in the image of f ”, i.e.

$$f : [k] \rightarrow [n] \text{ has property } P_i \iff \forall j \in [k]: f(j) \neq i.$$

With this definition, the surjective functions are exactly those functions that have no bad property, i.e. we need to count $X \setminus (X_1 \cup \dots \cup X_n)$.

Count $N(S)$: We claim $N(S) = (n-|S|)^k$, for any $S \subseteq [n]$. To see this, observe that f has all properties with indices from S if and only if $f(i) \notin S$ for all $i \in [k]$. In other words, f must be a function from $[k]$ to $[n] \setminus S$ and there are $(n-|S|)^k$ of those.

Apply PIE: Using Theorem 2.1, the number of surjections is therefore:

$$\begin{aligned} X \setminus (X_1 \cup \dots \cup X_n) &\stackrel{\text{PIE}}{=} \sum_{S \subseteq [n]} (-1)^{|S|} N(S) \\ &= \sum_{S \subseteq [n]} (-1)^{|S|} (n-|S|)^k \\ &= \sum_{i=0}^n (-1)^i \binom{n}{i} (n-i)^k. \end{aligned}$$

In the last step we used that $(-1)^{|S|} (n-|S|)^k$ only depends on the size of S and there are $\binom{n}{i}$ sets $S \subseteq [n]$ of size i . \square

Corollary 2.4. (i) Consider the case $n = k$. A function from $[n]$ to $[n]$ is a surjection if and only if it is a bijection. Since there are $n!$ bijections on $[n]$ (all permutations) we obtained the identity:

$$n! = \sum_{i=0}^n (-1)^i \binom{n}{i} (n-i)^n.$$

(ii) A surjection from $[k]$ to $[n]$ can be seen as a partition of $[k]$ into n non-empty distinguishable parts (the map assigns a part to each $i \in [n]$). Since the partitions of $[k]$ into n non-empty indistinguishable parts is counted by $s_n^{II}(k)$ and there are $n!$ ways to assign labels to the n parts, we obtain that the number of surjections is equal to $n!s_n^{II}(k)$. This proves the identity:

$$n!s_n^{II}(k) = \sum_{i=0}^n (-1)^i \binom{n}{i} (n-i)^k.$$

Theorem 2.5 (Derangements revisited). Recall that for $n \in \mathbb{N}$, the derangements D_n on n elements are permutations of $[n]$ without fixed points. We claim:

$$|D_n| = \sum_{i=0}^n (-1)^i \binom{n}{i} (n-i)!$$

Proof. **Define bad properties:** Let X be the set of all permutations of $[n]$. We define the “bad property” P_i that means “ π has a fixpoint i ”:

$$\pi \in X \text{ has property } P_i \iff \pi(i) = i, \quad (i \in [n]).$$

Derangements are exactly permutations that have none of these properties.

Count $N(S)$: We claim $N(S) = (n - |S|)!$ for any $S \subseteq [n]$.

Indeed, $\pi \in X$ has all properties with indices from S if and only if all $i \in S$ are fixed points of π . On the other elements, i.e. on $[n] \setminus S$, π may be an arbitrary bijection so there are $(n - |S|)!$ choices for π .

Apply PIE: Using Theorem 2.1, the number of derangements is therefore:

$$\begin{aligned} X \setminus (X_1 \cup \dots \cup X_n) &\stackrel{\text{PIE}}{=} \sum_{S \subseteq [n]} (-1)^{|S|} N(S) \\ &= \sum_{S \subseteq [n]} (-1)^{|S|} (n - |S|)! \\ &= \sum_{i=0}^n (-1)^i \binom{n}{i} (n-i)! \end{aligned}$$

In the last step we used that $(-1)^{|S|} (n - |S|)!$ only depends on the size of S and there are $\binom{n}{i}$ sets $S \subseteq [n]$ of size i . \square

Theorem 2.6 (Combinations of multisets). Consider a multiset M with types $1, \dots, m$ and repetition numbers r_1, \dots, r_m . Then the number of k -combinations of M is:

$$\sum_{S \subseteq [m]} (-1)^{|S|} \binom{m-1+k-\sum_{i \in S} (r_i+1)}{m-1}$$

where we define $\binom{a}{b} := 0$ for $a < b$.

Proof. Define bad properties: Let X be the set of k -combinations where we disregard the restrictions the repetition numbers impose, in other words X is the set of k -combinations of \widetilde{M} , where \widetilde{M} is the multiset with the same m types as in M but infinite supply of each type (i.e. “ $r_i = \infty$ ” for each $i \in [m]$).

Recall that by Theorem 1.9, $|X| = \binom{m-1+k}{m-1}$. Define the bad property P_i as:

$A \in X$ has property $P_i : \iff$ type i is repeated in A at least $r_i + 1$ times.

Then the k -combinations of M are exactly those k -combinations in X that have none of the bad properties.

Count N(S): We claim that for any $S \subseteq [m]$:

$$N(S) = \binom{m-1+k-\sum_{i \in S}(r_i+1)}{m-1}.$$

To see this, note that the repetition number of each type $i \in S$ is at least $r_i + 1$. This fixes $r_i + 1$ elements of the k -combination. In total $\sum_{i \in S}(r_i + 1)$ are fixed this way and what remains to be chosen is a $(k - \sum_{i \in S}(r_i + 1))$ -combination of \widetilde{M} . Again by Theorem 1.9 there are $\binom{m-1+k-\sum_{i \in S}(r_i+1)}{m-1}$ such combinations if $k - \sum_{i \in S}(r_i + 1) \geq 0$ and no such combinations otherwise.

Apply PIE: Using Theorem 2.1, the number of k -combinations of M is therefore:

$$\begin{aligned} X \setminus (X_1 \cup \dots \cup X_m) &\stackrel{\text{PIE}}{=} \sum_{S \subseteq [m]} (-1)^{|S|} N(S) \\ &= \sum_{S \subseteq [m]} (-1)^{|S|} \binom{m-1+k-\sum_{i \in S}(r_i+1)}{m-1} \quad \square \end{aligned}$$

In the special case where all the repetition numbers are equal, i.e. $r_1 = r_2 = \dots = r_m = r$, this can be simplified to:

$$\sum_{i=0}^m (-1)^i \binom{m}{i} \binom{m-1+k-(r+1) \cdot i}{m-1}.$$

Before we study a more advanced example, we prove a small result needed there:

Lemma. There are $\frac{2n}{2n-r} \binom{2n-r}{r}$ binary sequences (with letters 0 and 1) such that:

- The sequence has length $2n$ and exactly r copies of 1 and $2n - r$ copies of 0.
- No two copies of 1 are adjacent. Here the first and the last position of the sequence count as adjacent (the sequence is cyclic).

For example, if $n = 3$ and $r = 2$ there are 9 such sequences, namely:

101000, 100100, 100010, 010100, 010010, 010001, 001010, 001001, 000101.

Proof. Since no two copies of 1 may be adjacent, we know that after each 1 there must be a 0. So we can just imagine that one 0 is already “glued” to every 1 and we are actually building a sequence with $2n - 2r$ copies of 0 and r copies of 10.

However, one copy of 10 might “spill” across the border, i.e. the 1 could be in position $2n$ and the 0 in position 1. We need to handle this case separately.

Case 1: The last position of the sequence is 1. Then the first position is 0 and the remaining positions contain $r - 1$ copies of 10 and $2n - 2r$ copies of 0, now without any further pitfalls. There are $\binom{2n-r-1}{r-1}$ ways to arrange them.

Case 2: The last position of the sequence is 0. Then our special character 10 does not spill across the border and the sequence is any ordered arrangement of r copies of 10 and $2n - 2r$ copies of 0. There are $\binom{2n-r}{r}$ of them.

In total we have:

$$\binom{2n-r-1}{r-1} + \binom{2n-r}{r} = \frac{r}{2n-r} \binom{2n-r}{r} + \binom{2n-r}{r} = \frac{2n}{2n-r} \binom{2n-r}{r}.$$

□

Example 2.7 (Problème des ménages). There are $n \geq 2$ married couples at a dinner party, a husband and a wife each, denoted by H_1, \dots, H_n and W_1, \dots, W_n . They should be seated on a round table with $2n$ (distinguishable) seats such that:

- Men and women alternate, i.e. no two men and no two women sit next to each other. Equivalently, the seats with even labels are used either exclusively by women or exclusively by men.
- All couples are separated, i.e. no husband sits next to his wife.

We want to count how many such assignments exist.

First count the number of ways to seat the women. They may be seated in the even or odd seats and once this is fixed, there are $n!$ ways to seat them, so $2 \cdot n!$ ways in total.

Assuming the women are already seated, we now count the number of ways the men can be added. It is easy to see that this number does not depend on how the women are seated so we can assume without loss of generality that wife W_i sits in seat $2i$ and the odd-numbered seats are still free. We count the number of ways the men can join using PIE.

Define bad properties: Let X be the set of all ways to seat the men without paying attention to the rule that they must not sit next to their wives. There are $|X| = n!$ ways to do it.

We define the bad property P_i that captures that the husband H_i sits next to his wife W_i , i.e. on seat $2i - 1$ or on seat $2i + 1$ (all seat numbers are taken modulo $2n$, in particular, $2n$ and 1 are adjacent).

The permitted arrangements are exactly those with none of the bad properties.

Count $N(S)$: This time, calculating $N(S)$ for arbitrary $S \subseteq [n]$ is tricky. Instead we calculate

$$N^*(r) := \sum_{\substack{S \subseteq [n] \\ |S|=r}} N(S), \quad \text{for } 0 \leq r \leq n.$$

which will be just as helpful. The meaning of this number is a bit subtle: $N^*(r)$ is the number of pairs (π, S) where $S \subseteq [n]$ is of size r and π is a seating plan such that (wife W_i sits at $2i$ and) the couples with indices in S are not separated. To better count these pairs, define a map f :

Under f , the pair (π, S) is mapped to a binary sequence with $2n$ characters and exactly $|S|$ copies of 1: For $i \in S$ (remember that H_i and W_i will be assigned adjacent places in π) a one should be put in the position of the husband H_i – if he sits left of his wife – or in the position of W_i in π , if she sits left of her husband.

It is time for an example. Consider the pair (π, S) with $S = \{2, 5\}$ and the seating

$$\pi = (\overset{1}{H_5} \overset{2}{W_1} \overset{3}{H_2} \overset{4}{W_2} \overset{5}{H_1} \overset{6}{W_3} \overset{7}{H_4} \overset{8}{W_4} \overset{9}{H_3} \overset{10}{W_5}).$$

Note that the second and fifth couple are indeed not separated (W_5 and H_5 are adjacent because the table is circular). Also note that the fourth couple is also not separated, this is allowed. The mapping f discussed above would map this pair to the sequence 0010000001 since H_2 (in seat 3) and W_5 (in seat 10) sit left of their spouse. It is clear that f will never produce sequences with two adjacent ones: That would mean two adjacent people are both sitting left of their spouse: Which is impossible (assuming $n \geq 2$). However, any sequence with r copies of 1 and no two adjacent copies of 1 is the image of $(n-r)!$ pairs (π, S) : From the r copies of 1, the set S (of size r) can be reconstructed as can be the position of the husbands with indices in S . The $(n-r)$ other husbands can be distributed arbitrarily onto the $(n-r)$ remaining odd-numbered seats, there are $(n-r)!$ ways to do so. This proves, using the Lemma above to count the number of binary sequences of length n with r copies of 1 and no two adjacent copies of 1:

$$N^*(r) = (n-r)! \cdot \frac{2n}{2n-r} \binom{2n-r}{r}.$$

Apply PIE: Using Theorem 2.1, the number of valid ways to arrange the hus-

bands between the wives is:

$$\begin{aligned}
X \setminus (X_1 \cup \dots \cup X_n) &\stackrel{\text{PIE}}{=} \sum_{S \subseteq [n]} (-1)^{|S|} N(S) \\
&= \sum_{r=0}^n \sum_{\substack{S \subseteq [n] \\ |S|=r}} (-1)^{|S|} N(S) \\
&= \sum_{r=0}^n (-1)^r \sum_{\substack{S \subseteq [n] \\ |S|=r}} N(S) = \sum_{r=0}^n (-1)^r N^*(r) \\
&= \sum_{r=0}^n (-1)^r (n-r)! \cdot \frac{2n}{2n-r} \binom{2n-r}{r}.
\end{aligned}$$

Multiplying this with $2n!$, the number of ways to place the wives the total number of seatings is (for $n \geq 2$):

$$2n! \sum_{r=0}^n (-1)^r (n-r)! \cdot \frac{2n}{2n-r} \binom{2n-r}{r}.$$

2.1.2 Stronger Version of PIE

Theorem 2.1 can be strengthened as follows:

Theorem 2.8 (Stronger PIE). *Assume $f, g : 2^{[n]} \rightarrow \mathbb{R}$ are functions mapping subsets of $[n]$ to real numbers and g can be written in terms of f as:*

$$g(A) = \sum_{S \subseteq A} f(S) \quad (\text{for } A \subseteq [n]).$$

Then f can also be written in terms of g like this:

$$f(A) = \sum_{S \subseteq A} (-1)^{|A|-|S|} g(S) \quad (\text{for } A \subseteq [n]).$$

Before we proof Theorem 2.8, we first observe how it is indeed a generalization of Theorem 2.1. So assume we have the setting of Theorem 2.1, i.e. properties P_1, \dots, P_m with corresponding sets X_1, \dots, X_m .

Then for $S \subseteq [m]$ define $f(S)$ to be the number of elements having all properties with indices not in S and none of the properties with indices in S , i.e.

$$f(S) := \left| \bigcap_{i \in [m] \setminus S} X_i \setminus \bigcup_{i \in S} X_i \right|$$

Note that $f([m]) = |X| - |\bigcup_{i \in [m]} X_i|$ is precisely the number we want to count in Theorem 2.1.

The function g fulfilling the requirement of Theorem 2.8 is given as:

$$g(A) = \sum_{S \subseteq A} f(S) = \left| \bigcap_{i \in [m] \setminus A} X_i \right| = N([m] \setminus A)$$

To see the second “=”, note that $\sum_{S \subseteq A} f(S)$ counts an element x if and only if the set S of all properties that x does *not* have is a subset of A . This means x is counted if and only if it has all properties from $[m] \setminus A$.

We now apply Theorem 2.8 to obtain:

$$\begin{aligned} f([m]) &= \sum_{S \subseteq [m]} (-1)^{m-|S|} g(S) = \sum_{S \subseteq [m]} (-1)^{m-|S|} N([m] \setminus S) \\ &= \sum_{S \subseteq [m]} (-1)^{|S|} N(S). \end{aligned}$$

where in the last step we changed the order of summation, i.e. summed over $[m] \setminus S$ instead of S . This concludes the proof of (Thm. 2.8 \Rightarrow Thm. 2.1).

Theorem 2.8 establishes a similar result as Möbius Inversion which we consider later. In both cases there are two “linked” functions f and g where g is given in terms of f . The Theorems assert that f can then also be given in terms of g . The subset relation “ \subseteq ” will be replaced by the notion of divisibility “ $|$ ”. Both are order relations and, in fact, there are results even more general than both Theorem 2.8 and Möbius Inversions, dealing with general partial orders (but we will not consider them here).

We now proof Theorem 2.8.

Proof. Consider (for $A \subseteq [n]$) the term that is claimed to be equal to $f(A)$:

$$\sum_{S \subseteq A} (-1)^{|A|-|S|} g(S) = \sum_{S \subseteq A} (-1)^{|A|-|S|} \sum_{T \subseteq S} f(T) = \sum_{T \subseteq A} c_T f(T)$$

for appropriate c_T (to be determined!), that captures how often $f(T)$ is encountered (for $T \subseteq A$). Observe $c_A = 1$, since $f(A)$ is only encountered for $T = S = A$. For a proper subset $T \subset A$ (and $k := |A| - |T|$) we have:

$$c_T = \sum_{T \subseteq S \subseteq A} (-1)^{|A|-|S|} = \sum_{i=0}^k (-1)^{k-i} \binom{k}{i} = 0.$$

where in the second step we observed that picking a set between T and A is equivalent to picking a subset of $A \setminus T$. The last step is an identity we already saw.

This proves the claim. \square

2.2 Möbius Inversion Formula

Any positive number $n \in \mathbb{N}$ has a unique decomposition into primes, i.e. n can be written as $n = p_1^{a_1} \cdot p_2^{a_2} \cdot \dots \cdot p_k^{a_k}$ for primes p_1, \dots, p_k with multiplicities a_1, \dots, a_k . In other words, the factors of n are given as a multiset $M(n)$ with types p_1, \dots, p_k and repetition numbers a_1, \dots, a_k .

For example $360 = 2^3 \cdot 3^2 \cdot 5$ and $M(360) = \{2, 2, 2, 3, 3, 5\} = \{3 \cdot 2, 2 \cdot 3, 1 \cdot 5\}$ (which admittedly looks a bit quaint written that way).

We write $n|m$ if n divides m , meaning there is a number $a \in \mathbb{N}$ such that $a \cdot n = m$. In terms of multisets this means $M(n) \subseteq M(m)$.

The greatest common divisor (gcd) of m and n , is the number $k = \gcd(m, n)$, with $M(k) = M(m) \cap M(n)$. For example $\gcd(12, 90) = 6$ since $M(6) = \{2, 3\} = \{2, 2, 3\} \cap \{2, 3, 3, 5\} = M(12) \cap M(90)$.

Note that $M(1) = \emptyset$ (1 is the result of the empty product). If $\gcd(m, n) = 1$, then m and n are called *relatively prime* (or *coprime*).

We define the Euler's ϕ -function for $n \geq 2$ as:

$$\phi(n) = \#\{k \in \mathbb{N} \mid 1 \leq k \leq n, \gcd(n, k) = 1\}$$

For example, $\phi(12) = \#\{1, 5, 7, 11\} = 4$, $\phi(9) = \#\{1, 2, 4, 5, 7, 8\} = 6$. A formula to compute Euler's ϕ -function is given in the following Theorem.

Theorem 2.9. *If $n = p_1^{a_1} p_2^{a_2} \dots p_k^{a_k}$ then*

$$\phi(n) = n \prod_{i=1}^k \left(1 - \frac{1}{p_i}\right)$$

Consider for instance $n = 12 = 2^2 \cdot 3$. Then $\phi(12) = 12 \cdot (1 - \frac{1}{2}) \cdot (1 - \frac{1}{3}) = 12 \cdot \frac{1}{2} \cdot \frac{2}{3} = 4$ confirming what we counted by hand before.

Proof. In the proof we use the principle of inclusion and exclusion.

Define bad properties: As ground set we use $X = [n]$. We say a number $x \in X$ has the bad property P_i if x is divisible by p_i , i.e.:

$$x \in [n] \text{ has property } P_i \iff p_i | x. \quad (i \in [k]).$$

Then the numbers in $[n]$ that are relatively prime to n are exactly those that have no bad property.

Count N(S): For any $S \subseteq [k]$ we have $N(S) = \frac{n}{\prod_{i \in S} p_i}$.

This is because, if $x \in X$ is a multiple of all primes with indices from S , then x must be a multiple of their product $\prod_{i \in S} p_i$ (which is the least common multiple of those primes). Since this product divides n , there are no rounding issues and the number of numbers x between 1 and n that are multiples of $\prod_{i \in S} p_i$ is just given as the quotient.

Apply PIE: Using Theorem 2.1, we can write $\phi(n)$ as:

$$\phi(n) \stackrel{\text{PIE}}{=} \sum_{S \subseteq [k]} (-1)^{|S|} N(S) = n \cdot \sum_{S \subseteq [k]} \frac{(-1)^{|S|}}{\prod_{i \in S} p_i} = n \prod_{i \in [k]} \left(1 - \frac{1}{p_i}\right).$$

The last identity is best seen by multiplying out the right side: For each of the k factors we can either choose 1 or we can choose $-\frac{1}{p_i}$. The indices i where $-\frac{1}{p_i}$ was chosen are captured in the set S . For each $S \subseteq [k]$ we get exactly the term under the sum. \square

We now show a few more number theoretic results leading up to Möbius Inversions.

Theorem 2.10. $n = \sum_{d|n} \phi(d)$.

Proof. We claim that for a divisor d of n the sets $\{x \in [n] \mid \gcd(x, n) = d\}$ and $\{y \in [\frac{n}{d}] \mid \gcd(y, \frac{n}{d}) = 1\}$ have the same cardinality. Indeed, it is easy to see that $x \mapsto y := \frac{x}{d}$ is a bijection.

This means $\#\{x \in [n] \mid \gcd(x, n) = d\} = \phi(\frac{n}{d})$. Summing these identities for all $d \mid n$ yields:

$$n = \sum_{d \mid n} \phi(\frac{n}{d})$$

This is almost identical to the claim, the only thing left to do is to change the order of summation: Substitute d with $d' := \frac{n}{d}$ and note that d' divides n iff d divides n . \square

We now define the *Möbius Function* for $d \geq 1$ as:

$$\mu(d) := \begin{cases} 1 & d \text{ is the product of an even number of distinct primes} \\ -1 & d \text{ is the product of an odd number of distinct primes} \\ 0 & \text{otherwise} \end{cases}$$

For example, $\mu(15) = 1, \mu(7) = \mu(30) = -1, \mu(12) = 0$, since $15 = 3 \cdot 5$ is the product of two distinct primes, $7 = 7$ and $30 = 2 \cdot 3 \cdot 5$ are the product of an odd number of primes and $12 = 2 \cdot 2 \cdot 3$ is not the product of distinct primes: We need 2 twice. The numbers n with $\mu(n) \neq 0$ are also called *square free* since they do not have a square as a factor (12 is not square-free since it has 4 as a factor).

Note that $\mu(1) = 1$ since 1 is the product of 0 primes and 0 is even.

Lemma 2.11.

$$\sum_{d \mid n} \mu(d) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n \neq 1 \end{cases}$$

Proof. If $n = 1$ this is obvious since $d = 1$ is the only divisor of 1 and $\mu(1) = 1$.

For $n \geq 2$ we write $n = p_1^{a_1} \dots p_k^{a_k}$. Then we have:

$$\sum_{d \mid n} \mu(d) = \sum_{d \mid (p_1 p_2 \dots p_k)} \mu(d) = \sum_{D \subseteq [k]} (-1)^{|D|} = \sum_{i=0}^k \binom{k}{i} (-1)^i = 0.$$

In the first step remember that $\mu(d) = 0$ if d is not square-free. This means if d contains a prime factor more than once, it does not contribute to the sum. For the second step, realize that a divisor d of $p_1 \dots p_k$ is just given by choosing a subset D of those k primes and multiplying them. Then $\mu(d)$ will be 1 if an even number of primes were chosen and -1 otherwise. The last two steps are identities we already encountered earlier. \square

Corollary 2.12.

$$\frac{\phi(n)}{n} = \sum_{d \mid n} \frac{\mu(d)}{d}.$$

Proof. We use an identity that came up in the proof of Theorem 2.9 and arguments similar to those from the last Theorem.

$$\phi(n) = n \prod_{i=1}^k (1 - \frac{1}{p_i}) = n \sum_{S \subseteq [k]} \frac{-1^{|S|}}{\prod_{i \in S} p_i} = n \sum_{d \mid p_1 \dots p_k} \frac{\mu(d)}{d} = n \sum_{d \mid n} \frac{\mu(d)}{d}. \quad \square$$

We now have all ingredients to prove the Möbius Inversion Formula:

Theorem 2.13 (Möbius Inversion).

Let $f, g : \mathbb{N} \rightarrow \mathbb{R}$ be functions satisfying $g(n) = \sum_{d|n} f(d)$. Then:

$$f(n) = \sum_{d|n} \mu(d)g\left(\frac{n}{d}\right).$$

Proof. We start by changing the order of summation ($d \rightarrow \frac{n}{d}$) and using the definition of g .

$$\begin{aligned} \sum_{d|n} \mu(d)g\left(\frac{n}{d}\right) &= \sum_{d|n} \mu\left(\frac{n}{d}\right)g(d) \\ &= \sum_{d|n} \mu\left(\frac{n}{d}\right) \sum_{d'|d} f(d') \\ &= \sum_{d'|n} c_{d'} \cdot f(d') \end{aligned}$$

Where $c_{d'}$ are numbers (to be determined!) that count how often $f(d')$ occurs as a summand. Note that $c_n = \mu(1) = 1$ since $f(n)$ occurs only for $d' = d = n$. For $d'|n$, $d' \neq n$ we have:

$$c_{d'} = \sum_{d'|d|n} \mu\left(\frac{n}{d}\right) = \sum_{m|\frac{n}{d'}} \mu(m) = 0.$$

Where we substituted the summation index $d \mapsto m := \frac{n}{d}$ and used Lemma 2.11 in the last step. This proves the claim. \square

Example 2.14 (Circular Sequences of 0's and 1's). Circular sequences are for example:

$$A = \begin{array}{ccccccc} & & 0 & 1 & & & \\ & & & & 0 & & \\ 1 & & & & & & \\ & & & 0 & & & \\ 1 & & & & & & \\ & 0 & 1 & & 0 & & \end{array}, \quad B = \begin{array}{ccccccc} & & 1 & 0 & & & \\ & & & & 1 & & \\ 0 & & & & & & \\ & & & 1 & & & \\ 0 & & & & & & \\ & 0 & 1 & & 0 & & \end{array}, \quad C = \begin{array}{ccccccc} & & 1 & 0 & & & \\ & & & & 0 & & \\ 0 & & & & & & \\ & & & 1 & & & \\ 0 & & & & & & \\ & 1 & 0 & & 0 & & \end{array}$$

Circular sequences are considered equal if they can be transformed into one another by rotation. We would write

$$A = 001011010 = 110100010 = B \neq C = 100100100.$$

Let N_n be the number of circular 0/1-sequences of length n and $M(d)$ the number of *aperiodic* circular sequences of length d where a sequence is called aperiodic if it cannot be written as several times a shorter sequence. For example, C is not aperiodic since it consists of three copies of “100” while A (and therefore B) are aperiodic.

For every circular sequence S of length n there is a unique way to describe it as “repetitions of S' ” where S' is an aperiodic circular sequence. This is not

trivial, but also not hard to see. Clearly, the length of S' must divide n . So we found:

$$N_n = \sum_{d|n} M(d). \quad (\star)$$

Consider Table 2 for an example with $n = 6$.

d	M(d)	corresponding sequences
1	0,1	000000,111111
2	01	010101
3	001,011	001001,011011
6	000001,000011,000101, 000111,001011,010011, 001111,010111,011111	000001,000011,000101, 000111,001011,010011, 001111,010111,011111

Table 2: Aperiodic sequences of all lengths that divide 6. There is a bijection between them and the circular sequences of length 6.

Another crucial observation is that the number of aperiodic linear sequence of length d is just $d \cdot M(d)$ (every rotation of an aperiodic circular sequence) and therefore the number of all linear sequences of length n can be written as “repetition of some aperiodic linear sequence” so as:

$$2^n = \sum_{d|n} d \cdot M(d). \quad (\Delta)$$

Now define $g(n) = 2^n$ and $f(d) = d \cdot M(d)$. These choices of f and g fulfill the requirement of the Möbius Inversion Formula so we obtain:

$$\begin{aligned}
f(n) &\stackrel{2.13}{=} \sum_{d|n} \mu(d) g\left(\frac{n}{d}\right) \\
&\Leftrightarrow nM(n) = \sum_{d|n} \mu(d) 2^{\frac{n}{d}} \\
\text{so } N_n &\stackrel{(\star)}{=} \sum_{d|n} M(d) = \sum_{d|n} \frac{1}{d} \sum_{l|d} \mu(l) 2^{\frac{d}{l}} = \sum_{d|n} \frac{1}{d} \sum_{l|d} \mu\left(\frac{d}{l}\right) 2^l \\
&= \sum_{l|n} \sum_{l|d|n} \frac{1}{d} \mu\left(\frac{d}{l}\right) 2^l = \sum_{l|n} \sum_{1|\frac{d}{l}|\frac{n}{l}} \frac{1}{d} \mu\left(\frac{d}{l}\right) 2^l = \sum_{l|n} \sum_{\substack{k|\frac{n}{l} \\ (d:=k \cdot l)}} \frac{1}{k \cdot l} \mu(k) 2^l \\
&= \sum_{l|n} \frac{2^l}{l} \sum_{k|\frac{n}{l}} \frac{\mu(k)}{k} \stackrel{2.12}{=} \sum_{l|n} \frac{2^l}{l} \frac{\phi\left(\frac{n}{l}\right)}{\frac{n}{l}} = \sum_{l|n} \frac{2^l}{n} \phi\left(\frac{n}{l}\right).
\end{aligned}$$

While not overwhelmingly pretty, at least our result is an explicit formula only involving one sum and Euler’s ϕ -function. This is as good as it gets.

3 Generating Functions

In the following we consider sequences $(a_n)_{n \in \mathbb{N}} = a_0, a_1, a_2, \dots$ of non-negative numbers. Typically, a_k is the number of discrete structures of a certain type and “size” k .

The *generating function* for $(a_n)_{n \in \mathbb{N}}$ is given as $F(x) = \sum_{n=0}^{\infty} a_n x^n$. Despite the name, a generating function should not be thought of as a thing you plug values into: It is not meaningful to compute something like $F(5)$: In fact, these values are often not well-defined because the sum would be divergent. We care about the thing as a whole: You can either think of generating functions as functions that are well-defined within their radius of convergence (some area close to zero) or you just think of the “ x ” in $F(x)$ as an abstract thing (not a placeholder for a number) which makes $F(x)$ an element of the ring of formal power series. In the following, we ignore all technicalities and boldly apply analytic methods as though a generating function were just a simple well-behaved function. And it just works. If you think this is all a bit arcane, try to look at it this way:

At first, a generating function is just a silly way to write a sequence (instead of $a_0 = 1, a_1 = 42, a_2 = 23, \dots$ you would write $A(x) = 1 + 42x + 23x^2 + \dots$). Some operations on the sequence have a natural correspondence: For example, shifting the sequence ($a_0 = 0, a_1 = 1, a_2 = 42, a_3 = 23 \dots$) corresponds to multiplying $A(x)$ by x . Some complicated operations on sequences suddenly become simple and natural in the world of generating functions where analytic tools are readily available.

When dealing with generating functions $F(x) = \sum_{n=0}^{\infty} a_n x^n$ and $G(x) = \sum_{n=0}^{\infty} b_n x^n$, we freely use the following operations:

- Differentiate F term-wise

$$F'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1} = \sum_{n=0}^{\infty} (n+1) a_{n+1} x^n$$

- Multiply $F(x)$ by a scalar $\lambda \in \mathbb{R}$ term-wise

$$\lambda F(x) = \sum_{n=0}^{\infty} \lambda a_n x^n.$$

- Add $F(x)$ and $G(x)$

$$F(x) + G(x) = \sum_{n=0}^{\infty} (a_n + b_n) x^n.$$

- Multiply $F(x)$ and $G(x)$

$$F(x) \cdot G(x) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^n a_k b_{n-k} \right) x^n.$$

We will later see that the coefficients of the product sometimes count meaningful things if a_n and b_n did (see Example 3.2 and 3.3).

Example 3.1 (Maclaurin series).

- (i) Consider $(a_n)_{n \in \mathbb{N}}$ with $a_n = 1$ for all $n \in \mathbb{N}$. The corresponding generating function $F(x) = 1 + x + x^2 + x^3 + \dots$ is called the *Maclaurin series*.

We claim $F(x) = \frac{1}{1-x}$, which looks a lot like the identity for an infinite geometric series. To verify this, just observe that the product of $F(x)$ and $(1-x)$ is one:

$$(1-x)F(x) = (1+x+x^2+x^3+\dots) - (x+x^2+x^3+\dots) = 1.$$

- (ii) Differentiating $F(x) = \frac{1}{1-x}$ with respect to x on both sides yields:

$$F'(x) = \sum_{n=0}^{\infty} (n+1)x^n = \frac{1}{(1-x)^2}.$$

So we found the generating function for $(b_n)_{n \in \mathbb{N}}$ with $b_n = n+1$.

- (iii) Substituting $x = -y$ in the Maclaurin series gives the generating function of the alternating sequence (i.e. $a_n = (-1)^n$):

$$\frac{1}{1+y} = \sum_{n=0}^{\infty} (-1)^n y^n.$$

How do we count using Generating Functions?

Example 3.2. Let $a_n^{(k)}$ be the number of arrangements of n unlabeled balls in k labeled boxes and at least 1 ball per box. For each k , this gives a generating function:

$$B^{(k)}(x) = \sum_{n=0}^{\infty} a_n^{(k)} x^n.$$

Considering the case with only one box, we have

$$a_n^{(1)} = \begin{cases} 0 & \text{for } n = 0 \\ 1 & \text{for } n \geq 1. \end{cases}$$

So $B^{(1)}(x)$ is just the Maclaurin Series from 3.1(i) shifted by one position (i.e. multiplied by x), meaning:

$$B^{(1)}(x) = 0 + x + x^2 + x^3 + \dots = x \cdot \sum_{n=0}^{\infty} x^n = \frac{x}{1-x}.$$

The key observation we make now is that multiplying two generating functions has a meaningful correspondence in our balls-in-boxes setting:

For two numbers of boxes s and t , we have the identity:

$$a_n^{(s+t)} = \sum_{l=0}^n a_l^{(s)} a_{n-l}^{(t)}$$

This is merely the observation that arrangements of n balls in $s+t$ boxes are given by arranging some l balls in the first s boxes and the remaining $n-l$ balls

in the remaining t boxes. Looking at the right side of the equation, note that these numbers are exactly the coefficients of the product $B^{(s)}(x) \cdot B^{(t)}(x)$ (check this!). So we obtained:

$$B^{(s+t)}(x) = B^{(s)}(x) \cdot B^{(t)}(x)$$

and therefore

$$B^{(k)}(x) = \left(B^{(1)}(x)\right)^k = \left(\frac{x}{1-x}\right)^k.$$

However, we want to write $B^{(k)}(x)$ in the form $\sum_n a_n^{(k)} x^n$ to see the coefficients $a_n^{(k)}$. To do this, note that deriving $k-1$ times the term $(1-x)^{-1}$ yields $(k-1)!(1-x)^{-k}$. Using this we obtain:

$$\begin{aligned} B^{(k)}(x) &= \frac{x^k}{(1-x)^k} = \frac{x^k}{(k-1)!} \cdot \left(\frac{d^{k-1}}{dx^{k-1}} \frac{1}{1-x}\right) = \frac{x^k}{(k-1)!} \left(\frac{d^{k-1}}{dx^{k-1}} \sum_{n=0}^{\infty} x^n\right) \\ &= \frac{x^k}{(k-1)!} \sum_{n=k-1}^{\infty} n \cdot (n-1) \cdot \dots \cdot (n-k+2) x^{n-k+1} \\ &= \sum_{n=k-1}^{\infty} \frac{n!}{(n-k+1)!(k-1)!} x^{n+1} \\ &= \sum_{n=k-1}^{\infty} \binom{n}{k-1} x^{n+1} = \sum_{n=0}^{\infty} \binom{n-1}{k-1} x^n \end{aligned}$$

where in the last step use that $\binom{a}{b} = 0$ if $a < b$ and additionally make an index shift of 1.

So we found a new proof that the number of arrangements of n unlabeled balls in k labeled boxes and at least one ball per box is $a_n^{(k)} = \binom{n-1}{k-1}$.

Example 3.3. We want to count integer solutions for $a+b+c=n$ with a non-negative *even* integer a , a non-negative integer b and $c \in \{0, 1, 2\}$. Equivalently, we can think of this as counting arrangements of n unlabeled balls in boxes labeled a , b and c where box a should receive an even number of balls and c at most 2 balls.

We first consider the more simple situations where only one of the variables/boxes exists:

- Solutions to $a = n$ with even a . Clearly, there is a unique solution for even n and no solution for odd n . The corresponding generating function is:

$$A(x) = \sum_{n=0}^{\infty} x^{2n} = \sum_{n=0}^{\infty} (x^2)^n = \frac{1}{1-x^2}.$$

- Solutions to $b = n$ where b is an integer. Clearly, there is exactly one solution for each n . The corresponding generating function is the Maclaurin series:

$$B(x) = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$

- Solutions to $c = n$ where $c \in \{0, 1, 2\}$. Clearly, there is exactly one solution if $n \in \{0, 1, 2\}$ and no solution otherwise. The corresponding generating function is:

$$C(x) = 1 + x + x^2.$$

With the same argument as in the previous example, multiplying the generating functions yields the generating function of the sequence we are interested in, so we calculate:

$$F(x) = A(x) \cdot B(x) \cdot C(x) = \frac{(1 + x + x^2)}{(1 - x^2)(1 - x)} = \frac{(1 + x + x^2)}{(1 + x)(1 - x)^2}.$$

We would like to write this as a linear combination of generating function we understand well, so we search for real numbers R, S, T with:

$$F(x) = \frac{(1 + x + x^2)}{(1 + x)(1 - x)^2} = \frac{R}{1 + x} + \frac{S}{1 - x} + \frac{T}{(1 - x)^2}$$

Multiplying by the denominators yields:

$$(1 + x + x^2) = R(1 - x)^2 + S(1 - x^2) + T(1 + x)$$

We compare the coefficients in front of 1, x and x^2 and get the equations $1 = R + S + T$, $1 = -2R + T$, $1 = R - S$. This system has the unique solution $R = \frac{1}{4}$, $S = -\frac{3}{4}$, $T = \frac{3}{2}$.

Using this and the identities we obtained in Example 3.1 yields:

$$F(x) = \frac{1}{4} \frac{1}{1 + x} - \frac{3}{4} \frac{1}{1 - x} + \frac{3}{2} \frac{1}{(1 - x)^2} = \frac{1}{4} \sum_{n=0}^{\infty} (-1)^n x^n - \frac{3}{4} \sum_{n=0}^{\infty} x^n + \frac{3}{2} \sum_{n=0}^{\infty} (n+1) x^n.$$

So the coefficient of x^n , and therefore the number of solutions to $a + b + c = n$ we want to count is:

$$\frac{(-1)^n}{4} - \frac{3}{4} + \frac{3(n+1)}{2}.$$

Example 3.5. We want to count the number a_n of well-formed parenthesis expressions with n pairs of parenthesis. For example $((()))()$ is a well-formed expression with 4 pairs of parenthesis but $)()(($ is not. Formally, a permutation of the multiset $\{n \cdot “(”, n \cdot “)”\}$ is well-formed if reading it from left to right and counting “+1” for every “(” and “-1” for every “)” will never yield a negative number at any time (no prefix contains more closing than opening parenthesis).

Every well-formed expression with $n \geq 1$ pairs of parenthesis starts with “(” and there is a unique matching “)” such that the sequence in between and the sequence after is a well-formed (possibly empty) expression. For example:

$$()()() \quad (())() \quad ()(())()$$

In other words, a well-formed expression with n pairs of parenthesis is obtained by putting a well-formed expression with k pairs in between “(” and “)” and then append a well-formed expression with $n - k - 1$ pairs of parenthesis. This gives the equation:

$$a_n = \sum_{k=0}^{n-1} a_k a_{n-k-1}$$

So if $F(x)$ is the generating function belonging to a_n , then we know:

$$\begin{aligned} F(x) &= \sum_{n=0}^{\infty} a_n x^n = 1 + \sum_{n=1}^{\infty} \left(\sum_{k=0}^{n-1} a_k a_{n-k-1} \right) x^n = 1 + \sum_{n=0}^{\infty} \left(\sum_{k=0}^n a_k a_{n-k} \right) x^{n+1} \\ &= 1 + x \cdot \sum_{n=0}^{\infty} \left(\sum_{k=0}^n a_k a_{n-k} \right) x^n = 1 + x \cdot F(x)^2. \end{aligned}$$

And with analytic methods we can find a solution to this as:

$$F(x) = \frac{1 - \sqrt{1 - 4x}}{2x}.$$

So far it is unclear what this means for the coefficients of $F(x)$ (and therefore for the number of well-formed expressions). It seems we need additional tools to understand the power series of $\sqrt{1 - 4x}$.

3.1 Newton's Binomial Theorem

Recall the following special case of the Multinomial Theorem (Theorem 1.6).

$$(1 + x)^n = \sum_{k=0}^n \binom{n}{k} x^k = \sum_{k=0}^{\infty} \binom{n}{k} x^k, \quad \forall n \in \mathbb{N}.$$

Where, as usual, $\binom{n}{k} = 0$ for $k > n$. This shows that $(1 + x)^n$ is the generating function for the series $(a_k)_{k \in \mathbb{N}}$ with $a_k = \binom{n}{k}$.

We extend this result from natural numbers $n \in \mathbb{N}$ to any real number $n \in \mathbb{R}$. To this end we first extend the definition of binomial coefficients.

Definition 3.7 (Binomial Coefficients for Real Numbers). Recall that for integers $n, k \in \mathbb{N}$ we have:

$$\binom{n}{k} = \frac{|P(n, k)|}{k!} = \frac{n(n-1) \cdots (n-k+1)}{k!} = \frac{n!}{k!(n-k)!}$$

It does not make sense to talk about permutations of sets of size $n \in \mathbb{R}$ and it is unclear what $n!$ should be, but the formula in between is well-defined for general $n \in \mathbb{R}$. With this in mind, we define $p(n, k) := n \cdot (n-1) \cdots (n-k+1)$ and $\binom{n}{k} := \frac{p(n, k)}{k!}$. Note that the new definition matches the old one if n is integer.

We can now talk about numbers such as “ $-7/2$ choose 5” by which we mean:

$$\binom{-7/2}{5} = \frac{\frac{-7}{2} \cdot \frac{-9}{2} \cdot \frac{-11}{2} \cdot \frac{-13}{2} \cdot \frac{-15}{2}}{5!} = -\frac{9009}{256}.$$

Note that for $n \in \mathbb{R}$ we have $p(n, 0) = 1$ and for $k \geq 1$ the recursions:

$$\begin{aligned} p(n, k) &= (n - k + 1) \cdot p(n, k - 1) \\ &= n \cdot p(n - 1, k - 1). \end{aligned} \tag{*}$$

Given our extended definition of binomial coefficients, we can state the following Theorem (but will omit the proof).

Theorem 3.8 (Newton's Binomial Theorem). *For all non-zero $n \in \mathbb{R}$ we have:*

$$(1+x)^n = \sum_{k=0}^{\infty} \binom{n}{k} x^k.$$

Setting $n = 1/2$ yields an identity for $\sqrt{1+x}$ that we require to proceed in Example 3.5. But before we can use it, we need to better understand the coefficients of the form $\binom{1/2}{\cdot}$.

Lemma 3.9. *For any integer $n \geq 1$ we have*

$$\binom{1/2}{n} = (-1)^{n+1} \binom{2n-2}{n-1} \frac{1}{2^{2n-1}} \cdot \frac{1}{n}.$$

Proof. We do induction on n . For $n = 1$:

$$\binom{1/2}{1} = \frac{\frac{1}{2}}{1!} = \frac{1}{2} = 1 \cdot 1 \cdot \frac{1}{2} \cdot 1$$

For the induction step ($n \rightarrow n+1$) we use recursion (\star) from above:

$$\begin{aligned} \binom{1/2}{n+1} &= \frac{p(1/2, n+1)}{(n+1)!} = \frac{(1/2 - (n+1) + 1)p(1/2, n)}{(n+1) \cdot n!} = -\frac{n-1/2}{n+1} \binom{1/2}{n} \\ &\stackrel{\text{IH}}{=} -\frac{n-1/2}{n+1} (-1)^{n+1} \binom{2n-2}{n-1} \frac{1}{2^{2n-1}} \cdot \frac{1}{n} \\ &= \frac{2n}{2n} \cdot \frac{2n-1}{2n} (-1)^{n+2} \binom{2n-2}{n-1} \frac{1}{2^{2n-1}} \cdot \frac{1}{n+1} \\ &= (-1)^{n+2} \underbrace{\frac{(2n-2)!(2n-1)(2n)}{(n-1)!(n-1) \cdot n \cdot n}}_{\binom{2n}{n}} \frac{1}{2^{2n+1}} \cdot \frac{1}{n+1}. \quad \square \end{aligned}$$

Proposition 3.10. *Using the last Theorem and Lemma we obtain:*

$$\sqrt{1+x} \stackrel{3.8}{=} \sum_{n=0}^{\infty} \binom{1/2}{n} x^n \stackrel{3.9}{=} 1 + \sum_{n=1}^{\infty} -2 \binom{2n-2}{n-1} (-1)^n \frac{1}{2^{2n}} \cdot \frac{1}{n} \cdot x^n.$$

Example 3.5 (Continued). Using the proposition we are now able to find the coefficients of the generating function from Example 3.5 above, i.e. the number of well-formed parenthesis expressions.

$$\begin{aligned} F(x) &= \frac{1 - \sqrt{1-4x}}{2x} \stackrel{3.10}{=} \frac{1}{2x} \sum_{n=1}^{\infty} 2 \binom{2n-2}{n-1} (-1)^n \frac{1}{2^{2n}} \frac{1}{n} (-4x)^n \\ &= \frac{1}{x} \sum_{n=1}^{\infty} \binom{2n-2}{n-1} \frac{1}{n} x^n = \sum_{n=0}^{\infty} \binom{2n}{n} \frac{1}{n+1} x^n. \end{aligned}$$

The numbers $C_n := \binom{2n}{n} \frac{1}{n+1}$ are called *Catalan Numbers*. They do not only count well-formed parenthesis expressions but occur in other situations as well.

3.2 Exponential Generating Functions

Until now we mapped sequences to functions like this:

$$(a_n)_{n \in \mathbb{N}} \mapsto \sum_{n=0}^{\infty} a_n x^n \in \mathbb{R}[x]$$

In other words, we used the coefficients a_n to obtain a linear combination of the basis $\{x^n\}_{n \in \mathbb{N}}$ of $\mathbb{R}[x]$.

Our choice of a basis was useful in the cases we considered, but we can consider other bases as well that will be useful in other situations. The following sets are all bases of $\mathbb{R}[x]$:

$$\left\{ p(x, n) \right\}_{n \in \mathbb{N}} \quad \left\{ \frac{x^n}{n!} \right\}_{n \in \mathbb{N}} \quad \left\{ e^{-x} \frac{x^n}{n!} \right\}_{n \in \mathbb{N}} \quad \left\{ \frac{1}{n^x} \right\}_{n \in \mathbb{N}}$$

In the last case we get from any sequence $(a_n)_{n \in \mathbb{N}}$ a corresponding *Dirichlet series* $\sum_{n=0}^{\infty} \frac{a_n}{n^s}$. Such series are important in algebraic number theory (in that setting the variable is typically called s instead of x). As an example, we state the following result (without proof).

Theorem 3.11 (Euler Product). *The Dirichlet series for $(\mu(n))_{n \in \mathbb{N}}$ satisfies*

$$\sum_{n=0}^{\infty} \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)} = \prod_{p \text{ prime}} (1 - p^{-s})$$

where $\zeta(s)$ is the Riemann zeta function $\zeta(s) := \sum_{n=0}^{\infty} \frac{1}{n^s}$.

We will not examine Dirichlet series further, dealing with *exponential generating functions* instead. In contrast to *ordinary generating functions* (the kind we considered before) the basis is not $\{x^n\}_{n \in \mathbb{N}}$ but $\{\frac{x^n}{n!}\}_{n \in \mathbb{N}}$ and instead of counting arrangements of unlabeled objects (unlabeled balls in labeled boxes, solutions to $a_1 + a_2 + \dots + a_l = n$ with constraints, indistinguishable parenthesis in an ordered string), exponential generating functions are useful to count arrangements of *labeled* objects (permutations, derangements, partitions, ...) as we will see shortly.

Before we get started, note the exponential generating functions $A(x)$ and $B(x)$ of $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$ with $a_n = 1$ and $b_n = n!$ ($n \in \mathbb{N}$) are

$$A(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x$$

$$B(x) = \sum_{n=0}^{\infty} n! \frac{x^n}{n!} = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$

The following two observations show the connection of exponential generating functions to arrangement of labeled objects.

Observation 3.12. Say the numbers $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$ count arrangements of type A and B , respectively, using n labeled objects.

Assume arrangements of type C with n objects are obtained by a unique split of the n objects into two sets and then forming an arrangement of type A with the first set and an arrangement of type B with the second.

Then c_n , the number of arrangements of type C and size n , is given as

$$c_n = \sum_{k=0}^n \binom{n}{k} a_k \cdot b_{n-k}.$$

Crucially, the exponential generating functions $A(x), B(x), C(x)$ of the three sequences reflect this relationship as $C(x) = A(x) \cdot B(x)$, which is easy to verify:

$$\begin{aligned} A(x) \cdot B(x) &= \left(\sum_{n=0}^{\infty} a_n \frac{x^n}{n!} \right) \cdot \left(\sum_{n=0}^{\infty} b_n \frac{x^n}{n!} \right) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^n \frac{a_k}{k!} \frac{b_{n-k}}{(n-k)!} \right) x^n \\ &= \sum_{n=0}^{\infty} \left(\sum_{k=0}^n \binom{n}{k} a_k \cdot b_{n-k} \right) \frac{x^n}{n!} = C(x). \end{aligned}$$

Observation 3.13. An index shift by one in the sequence corresponds to a derivation, meaning

$$F(x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{n!} \xrightarrow{\text{derive}} F'(x) = \sum_{n=1}^{\infty} a_n \frac{x^{n-1}}{(n-1)!} = \sum_{n=0}^{\infty} a_{n+1} \frac{x^n}{n!}.$$

We already encountered Stirling numbers before. We distinguish three kinds:

- The *unsigned Stirling numbers of the first kind* $s_k^I(n)$ fulfill the recursion

$$s_k^I(n) = (n-1)s_k^I(n-1) + s_{k-1}^I(n-1)$$

and count the number of n -permutations of $[n]$ with k cycles.

- The *signed Stirling numbers of the first kind* are simply

$$\bar{s}_k^I(n) = (-1)^{n-k} s_k^I(n).$$

- The *Stirling numbers of the second kind* $s_k^{II}(n)$ fulfill

$$s_k^{II}(n) = k s_k^{II}(n-1) + s_{k-1}^{II}(n-1)$$

and count the number of partitions of $[n]$ into k non-empty parts.

Theorem 3.14. For $k \in \mathbb{N}$, we have the following identity for the exponential generating function $F_k(x)$ of the Stirling numbers $(s_k^{II}(n))_{n \in \mathbb{N}}$:

$$F_k(x) = \sum_{n=0}^{\infty} s_k^{II}(n) \frac{x^n}{n!} = \frac{1}{k!} (e^x - 1)^k.$$

Proof. We proceed by induction on k . For $k = 1$ we have:

$$s_k^{II}(n) = \begin{cases} 1 & \text{for } n \geq 1 \\ 0 & \text{for } n = 0 \end{cases}$$

and the claim follows from the identity $\sum_{n=1}^{\infty} \frac{x^n}{n!} = e^x - 1$.

If $k \geq 2$, we use the recursion of Stirling numbers from above:

$$s_k^{II}(n+1) = k s_k^{II}(n) + s_{k-1}^{II}(n).$$

Taking the generating functions for each term (remember that derivations correspond to index shifts!) and then applying the induction hypothesis yields:

$$F_k'(x) = k \cdot F_k(x) + F_{k-1}(x) \stackrel{\text{IH}}{=} k F_k(x) + \frac{1}{(k-1)!} (e^x - 1)^{k-1}.$$

This differential equation has the solution

$$F_k(x) = \frac{1}{k!} (e^x - 1)^k$$

as claimed, where we used $F_k(0) = 0$. □

Note that solutions to differential equations may be hard to find but are easy to verify. Finding the solutions is not the topic of this lecture. It is perfectly alright if you use computers (or techniques from other lectures).

To later obtain the exponential generating function for Stirling numbers of the first kind, we first prove two identities:

Lemma 3.15.

$$(i) \quad \sum_{k=0}^n s_k^I(n) x^k = p(x+n-1, n).$$

$$(ii) \quad \sum_{k=0}^n \bar{s}_k^I(n) x^k = p(x, n).$$

Proof. Note that $p(x+n-1, n)$ is a polynomial of degree n with variable x so there are constants $b_{n,k}$ such that $p(x+n-1, n) = \sum_{k=0}^n b_{n,k} x^k$. They fulfill the following equation:

$$\begin{aligned} \sum_{k=0}^n b_{n,k} x^k &= p(x+n-1, n) = (x+n-1) \cdot p(x+n-2, n-1) \\ &= (x+n-1) \sum_{k=0}^{n-1} b_{n-1,k} x^k = \sum_{k=1}^n b_{n-1,k-1} x^k + (n-1) \sum_{k=0}^{n-1} b_{n-1,k} x^k. \end{aligned}$$

Comparing the coefficients of x_k we get:

$$b_{n,k} = b_{n-1,k-1} + (n-1)b_{n-1,k}$$

where we define $b_{n,k} = 0$ for $k > n$ or $k < 0$.

So the numbers $b_{n,k}$ fulfill the same recursion as $s_k^I(n)$ and since the starting values $s_0^I(0) = b_{0,0} = 1$ match as well, this proves $b_{n,k} = s_k^I(n)$ and thus (i).

For (ii), plug $-x$ into the equation above:

$$\begin{aligned} \sum_{k=0}^n (-1)^k s_k^I(n) x^k &= p(-x + n - 1, n) = \prod_{k=1}^n (-x + n - k) \\ &= (-1)^n \prod_{k=1}^n (x - n + k) = (-1)^n p(x, n). \end{aligned}$$

Multiplying by $(-1)^n$ gives the desired result:

$$\sum_{k=0}^n \underbrace{(-1)^{n-k} s_k^I(n)}_{\bar{s}_k^I(n)} x^k = p(x, n). \quad \square$$

Theorem 3.16. *The exponential generating function of $(\bar{s}_k^I(n))_{n \in \mathbb{N}}$ fulfills:*

$$\sum_{n=0}^{\infty} \bar{s}_k^I(n) \frac{x^n}{n!} = \frac{1}{k!} (\log(1+x))^k.$$

Proof. One way to expand $(1+x)^z$ is:

$$(1+x)^z = e^{z \log(1+x)} = \sum_{k=0}^{\infty} \frac{1}{k!} (\log(1+x))^k z^k$$

On the other hand, we can also expand $(1+x)^z$ using Newton's Binomial Theorem and the last Lemma:

$$\begin{aligned} (1+x)^z &\stackrel{3.8}{=} \sum_{n=0}^{\infty} \binom{z}{n} x^n = \sum_{n=0}^{\infty} \frac{p(z, n)}{n!} x^n \stackrel{3.15}{=} \sum_{n=0}^{\infty} \frac{x^n}{n!} \sum_{k=0}^n \bar{s}_k^I(n) z^k \\ &\stackrel{\uparrow}{=} \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \bar{s}_k^I(n) \frac{x^n}{n!} z^k = \sum_{k=0}^{\infty} \left(\sum_{n=0}^{\infty} \bar{s}_k^I(n) \frac{x^n}{n!} \right) z^k. \end{aligned}$$

$\begin{smallmatrix} \bar{s}_k^I(n)=0 \\ \text{for } k > n \end{smallmatrix}$

We have written $(1+x)^z$ in two ways and the coefficients of z^k in both representations must match. This proves the claim. \square

3.3 Recurrence Relations

We have already seen (in the case of Stirling numbers) that relations between numbers in a sequence may fully characterize the sequence (if some starting value is given). A sequence $(a_n)_{n \in \mathbb{N}}$ is defined *recursively* if

$$a_n = f(a_{n-1}, a_{n-2}, \dots, a_{n-k})$$

for some function f , a positive integer k , and all $n \geq k$. This identity $a_n = f(a_{n-1}, a_{n-2}, \dots, a_{n-k})$ is called a *recurrence* or a *recurrence relation* or a *recursion* of order k . If a_0, a_1, \dots, a_{k-1} are known, they are called *initial values* or *initial conditions*. The sequence $(a_n)_{n \in \mathbb{N}}$ is then called a *solution* of the recursion f with initial values a_0, \dots, a_{k-1} .

In the following we examine special cases of such recurrence relations and show how the sequence can be derived from them.

Example 3.17 (Fibonacci Numbers). Rabbits were first brought to Australia by settlers in 1788. Assume for simplicity, the first pair of young rabbits (male and female) arrives in January (at time $n = 1$). A month later ($n = 2$) this pair of rabbits reaches adulthood. Another month later ($n = 3$) they produced a new young pair of rabbits as offspring. In general, assume that in the course of a month every pair of young rabbits grows into adulthood and every pair of adult rabbits produces one pair of young rabbits as offspring. If F_n denotes the number of rabbits at time n , then we have $F_n = F_{n-1} + F_{n-2}$, since exactly the rabbits that already existed at time $n - 2$ will be adults at time $n - 1$ and produce offspring between time $n - 1$ and n . This gives sequence:

$$F_0 = 0, F_1 = 1, F_2 = 1, F_3 = 2, F_4 = 3, F_5 = 5, F_6 = 8, F_7 = 13, F_8 = 21, \dots$$

The sequence $(F_n)_{n \in \mathbb{N}}$ is called the *Fibonacci sequence* and its elements are the famous *Fibonacci numbers*. Rabbits in Australia are now a serious problem, causing substantial damage to crops and have been combated with ferrets, fences, poison, firearms and the myxoma virus.



Photos taken by Daniel Oines and Esdras Calderan

Figure 19: Spiraling patterns can be seen in many plants. The pine cone on the left has 8 spiral arms spiraling out counterclockwise and 13 spiral arms spiraling out clockwise. In the sunflower, different spiral patterns stand out depending on where you look. Close to the center, a counterclockwise spiral with 21 arms and a clockwise spiral with 34 arms can be seen. Closer to the border another counterclockwise spiral with 55 arms emerges. Fibonacci numbers everywhere! We're not making this up, zoom in and count for yourselves (or better yet, go outside and look at nature directly), the patterns are really there!

Fibonacci numbers also frequently occur in plants (see Figure 19). In case you do not know Vi Hart (seriously?!) you should definitely check out her [Youtube Channel](#) and watch her [videos on this topic](#).

Another example where Fibonacci numbers occur is tilings of checkerboards of size $n \times 2$ with dominoes as shown in Figure 20. Let c_n be the number of such tilings. Note that there are two possibilities to cover the rightmost column (see Figure 21: Either with a vertical domino, then a board of size $(n - 1) \times 2$

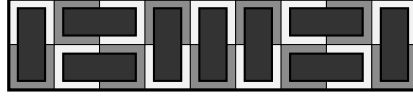


Figure 20: A tiling of the 9×2 board with dominoes.

remains to be tiled, or with horizontal dominoes, then a board of size $(n-2) \times 2$ remains to be tiled. This means $c_n = c_{n-1} + c_{n-2}$ and together with $c_1 = 1$ and $c_2 = 2$ (and setting $c_0 = 1$) this implies $c_n = F_{n+1}$.

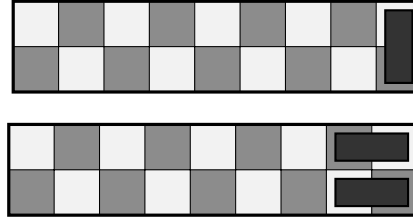


Figure 21: The last column can either be covered by a vertical domino or by horizontal dominoes.

Example 3.18 (Ternary n -strings without substring 20). Consider strings over the alphabet $\{0, 1, 2\}$ that do not contain the substring 20, for instance 0010221021 or 1111011102. Let t_n be the number of such strings of length n . We have $t_0 = 1$ (the empty string), $t_1 = 3$ (each string of length 1), and $t_2 = 8$ (all strings of length 2 except for 20).

In general, an n -string without 20 consists of an $(n-1)$ -string without 20 and an additional digit, which gives $3 \cdot t_{n-1}$ possibilities. However, this may produce strings that end in 20 (but have no 20 in other places): There are t_{n-2} of them.

This gives $t_n = 3 \cdot t_{n-1} - t_{n-2}$.

The Fibonacci sequence and the sequence $(t_n)_{n \in \mathbb{N}}$ from the last example have in common that they are solutions for some recurrence relation of order 2, i.e., their elements can be described in terms of the previous two elements. Moreover, each element could be expressed as the sum of multiples of previous elements. We now develop a general theory dealing with such recurrence relations.

Definition 3.19. A recurrence relation of the form

$$a_n = c_1(n)a_{n-1} + c_2(n)a_{n-2} + \cdots + c_k(n)a_{n-k} + g(n)$$

for functions $g : \mathbb{N} \rightarrow \mathbb{R}$ and $c_i : \mathbb{N} \rightarrow \mathbb{R}$, $i = 1, \dots, k$, is called a *linear* recurrence. In case $g \equiv 0$ (that is, $g(n) = 0$ for all $n \in \mathbb{N}$), the recurrence is called *linear homogeneous*, otherwise *linear non-homogeneous*.

In this chapter we shall treat the case when the coefficients c_i 's are constants, so

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k}.$$

We assume that $c_k \neq 0$, otherwise we treat the recursion as one of a smaller order.

The Fibonacci sequence fulfills a homogeneous linear recurrence relation with constant coefficients

$$F_n = F_{n-1} + F_{n-2}$$

where $k = 2, c_1 = 1, c_2 = 1, g \equiv 0$.

The same is true for t_n from Example 3.18

$$t_n = 3t_{n-1} - t_{n-2}$$

where this time $k = 2, c_1 = 3, c_2 = -1, g \equiv 0$.

3.3.1 Special Solution $a_n = x^n$ and the Characteristic Polynomial

Throughout this section assume we are given a linear homogeneous recurrence relation f of the form

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \dots + c_k a_{n-k}$$

of order k . If one replaces a_i with x^i and plugs it in the recursion, the following is obtained:

$$x^n = c_1 x^{n-1} + c_2 x^{n-2} + \dots + c_k x^{n-k}.$$

Dividing by x^{n-k} and bringing terms to one side of the equation gives:

$$x^k - c_1 x^{k-1} - c_2 x^{k-2} - \dots - c_k = 0.$$

Solving this equation allows to find some solutions. The *characteristic polynomial* of f is given by

$$p(x) = x^k - c_1 x^{k-1} - c_2 x^{k-2} - \dots - c_{k-1} x - c_k.$$

We see that a homogeneous linear recursion could be reconstructed from its characteristic polynomial by replacing x^i with a_{n-k+i} :

$$\begin{aligned} a_n - c_1 a_{n-1} - c_2 a_{n-2} - \dots - c_k a_{n-k} &= 0, \\ x^k - c_1 x^{k-1} - c_2 x^{k-2} - \dots - c_{k-1} x - c_k &= 0. \end{aligned}$$

Note further, that we can assume that all the roots of the characteristic polynomial are non-zero because $c_k \neq 0$.

In this section we shall present some methods to solve linear homogeneous recurrences and some specific non-homogeneous recurrences.

Lemma 3.20 (Linearity Property). *Let f be a linear homogeneous recurrence with constant coefficients. If $(b_n)_{n \in \mathbb{N}}$ is a solution of f with some initial values and $(d_n)_{n \in \mathbb{N}}$ is a solution of f with some initial values, then for any constants $\beta, \delta \in \mathbb{R}$, $(\beta b_n + \delta d_n)_{n \in \mathbb{N}}$ is a solution of f with some initial values.*

Proof. Let $a_n = f(a_{n-1}, \dots, a_{n-k}) = c_1 a_{n-1} + \dots + c_k a_{n-k}$. Thus $b_n = c_1 b_{n-1} + \dots + c_k b_{n-k}$ and $d_n = c_1 d_{n-1} + \dots + c_k d_{n-k}$ for all $n \geq k$. Multiplying the first equation by β , the second by δ and adding them gives $\beta b_n + \delta d_n = c_1(\beta b_{n-1} + \delta d_{n-1}) + \dots + c_k(\beta b_{n-k} + \delta d_{n-k})$. \square

Lemma 3.21 (Divisibility of Polynomials). *Let $p(x)$ be a polynomial. Then $p(x)$ is divisible by $(x - q)^m$, for a positive integer m and a nonzero constant q if and only if the following polynomials are divisible by $(x - q)$: $p_0(x) = p(x)$, $p_1(x) = xp'_0(x)$, $p_2(x) = xp'_1(x)$, \dots , $p_k(x) = xp'_{k-1}(x)$, \dots , $p_{m-1}(x) = xp'_{m-2}(x)$.*

Proof. Assume first that $p(x)$ is divisible by $(x-q)^m$. Then $p(x) = (x-q)^m t(x)$, where $t(x)$ is a polynomial. We use induction on k to show that $(x-q)^{m-k}$ divides p_k , $k = 0, \dots, m-1$. When $k = 0$, $(x-q)^{(m-0)}$ divides $p_0(x) = p(x)$. Assume that $(x-q)^{m-k+1}$ divides $p_{k-1}(x)$. Let us show that $(x-q)^{m-k}$ divides $p_k(x)$. We have then that $p_{k-1} = (x-q)^{m-k+1}s(x)$ for some polynomial s . Then $p_k(x) = xp'_{k-1}(x) = x((m-k+1)(x-q)^{m-k}s(x) + (x-q)^{m-k+1}s'(x))$, so p_k is divisible by $(x-q)^{m-k}$.

Now, assume that each of p_0, \dots, p_{m-1} is divisible by $x-q$. We shall show that $p_{m-k}(x)$ is divisible by $(x-q)^k$, $k = 0, \dots, m-1$, so in particular that $p_0(x)$ is divisible by $(x-q)^m$. We do this by induction on k . When $k = 1$, the statement holds by assumption. Now, assume that $p_{m-k}(x)$ is divisible by $(x-q)^k$, let us prove that $p_{m-k-1}(x)$ is divisible by $(x-q)^{k+1}$. Assume not, i.e., that $p_{m-k-1}(x) = (x-q)^\ell t(x)$, $t(q) \neq 0$, $\ell \leq k$. Then $p_{m-k}(x) = x(x-q)^{\ell-1}(\ell t(x) + t'(x)(x-q))$, so the highest power of $(x-q)$ dividing p_{m-k} is $\ell-1 < k$. We see that $p_{m-k}(x)$ is not divisible by $(x-q)^k$. \square

Lemma 3.22 (Fundamental Solution). *Let f be a linear homogeneous recurrence with constant coefficients. If q is a root of the characteristic polynomial of f , then $(q^n)_{n \in \mathbb{N}}$ is a solution of f for some initial values. Moreover, if q is a root of the characteristic polynomial of multiplicity s , $s > 1$, then $(n^i q^n)_{n \in \mathbb{N}}$, $0 \leq i \leq s-1$ is a solution of f .*

Proof. Let $a_n = f(a_{n-1}, \dots, a_{n-k}) = c_1 a_{n-1} + \dots + c_k a_{n-k}$ and $b_n = q^n$, where q is a root of the characteristic polynomial, so $q^k - c_1 q^{k-1} - \dots - c_k q^0 = 0$. Multiply the last equality by q^{n-k} to obtain $q^n - c_1 q^{n-1} - \dots - c_k q^{n-k} = 0$. So $(q^n)_{n \in \mathbb{N}}$ is a solution of f .

Now, assume that q has multiplicity s , i.e., the characteristic polynomial $p_1(x)$ has a form $p_1(x) = (x-q)^s p(x)$, where $p(x)$ is a polynomial. Observe that q is a root of each of the following polynomials: $p_2(x) = p_1(x)x^{n-k}$, $p'_2(x)$, $p_3(x) = xp'_2(x)$, $p'_3(x)$, $p_4(x) = xp'_3(x)$, etc. till $p_{s+2}(x)$. Indeed, the $(x-q)$ term remains after performing the product rule. So, in particular, for $i \leq s$

$$p_{i+2}(q) = 0. \quad (2)$$

We have

$$\begin{aligned} p_2(x) &= x^n - c_1 x^{n-1} - \dots - c_k x^{n-k}, \\ p'_2(x) &= nx^{n-1} - c_1(n-1)x^{n-2} - \dots - c_k(n-k)x^{n-k-1}, \\ p_3(x) &= nx^n - c_1(n-1)x^{n-1} - \dots - c_k(n-k)x^{n-k}, \\ p'_3(x) &= n^2 x^{n-1} - c_1(n-1)^2 x^{n-2} - \dots - c_k(n-k)^2 x^{n-k-1}, \\ p_4(x) &= n^2 x^n - c_1(n-1)^2 x^{n-1} - \dots - c_k(n-k)^2 x^{n-k}, \\ &\vdots \\ p_{i+2}(x) &= n^i x^n - c_1(n-1)^i x^{n-1} - \dots - c_k(n-k)^i x^{n-k}. \end{aligned}$$

We need to verify that for $b_n = n^i q^n$, $(b_n)_{n \in \mathbb{N}}$ gives a solution of the recurrence $a_n - c_1 a_{n-1} - \dots - c_k a_{n-k} = 0$. Let us plug b_n for a_n in the left hand side of

the equation:

$$\begin{aligned} b_n - c_1 b_{n-1} - \cdots - c_k b_{n-k} &= \\ n^i q^n - c_1 (n-1)^i q^{n-1} - \cdots - c_k (n-k)^i q^{n-i} &= \\ p_{i+2}(q) &= 0. \end{aligned}$$

Here the last equality holds by (2). \square

Theorem 3.23 (Existence and Uniqueness). *Let f be a linear homogeneous recursion with constant coefficients and $p(x)$ be its characteristic polynomial with roots q_1, \dots, q_r having multiplicities s_1, \dots, s_r , respectively. Then for any initial values, there is a solution $(a_n)_{n \in \mathbb{N}}$ of f written in the form*

$$a_n = (\star q_1^n + \star n q_1^n + \cdots + \star n^{s_1-1} q_1^n) + \cdots + (\star q_r^n + \star n q_r^n + \cdots + \star n^{s_r-1} q_r^n), \quad (3)$$

where \star 's are some constants. In particular, if all roots of the characteristic polynomial are distinct, $a_n = \star q_1^n + \star q_2^n + \cdots + \star q_r^n$.

Proof. From Lemma 3.22, all the listed terms are solutions of the recursion with some initial values. From Lemma 3.20, any of their linear combinations is a solution of the recursion with some initial values. It remains to verify that there are coefficients (\star 's) of the linear combination so that it is a solution of the recursion with the given initial values a_0, a_1, \dots, a_{k-1} . Denote the consecutive coefficient in (3) in front of $n^j q_i^n$ by $\gamma_{i,j}$, $i = 1, \dots, r$, $j = 0, \dots, s_i - 1$. That is, the $\gamma_{i,j}$'s are the unknowns solving equation (3) for all $n \in \mathbb{N}$, which can hence be written as

$$a_n = \sum_{i=1}^r \sum_{j=0}^{s_i-1} \gamma_{i,j} n^j q_i^n = \sum_{i=1}^r \left(\sum_{j=0}^{s_i-1} \gamma_{i,j} n^j \right) q_i^n.$$

Recall that a_0, a_1, \dots, a_{k-1} are the known initial values. Plugging $n = 0, n = 1, \dots, n = k-1$ we get:

$$\begin{aligned} a_0 &= (\gamma_{1,0} + 0 + \cdots + 0) + \cdots + (\gamma_{r,0} + 0 + \cdots + 0) \\ a_1 &= (\gamma_{1,0} + \gamma_{1,1} + \cdots + \gamma_{1,s_1-1}) q_1 + \\ &\quad \cdots + (\gamma_{r,0} + \cdots + \gamma_{r,s_r-1}) q_r \\ a_2 &= (\gamma_{1,0} + 2\gamma_{1,1} + \cdots + 2^{s_1-1} \gamma_{1,s_1-1}) q_1^2 + \\ &\quad \cdots + (\gamma_{r,0} + \cdots + 2^{s_r-1} \gamma_{r,s_r-1}) q_r^2 \\ &\vdots \\ a_{k-1} &= (\gamma_{1,0} + (k-1)\gamma_{1,1} + \cdots + (k-1)^{s_1-1} \gamma_{1,s_1-1}) q_1^{k-1} + \\ &\quad \cdots + (\gamma_{r,0} + \cdots + (k-1)^{s_r-1} \gamma_{r,s_r-1}) q_r^{k-1} \end{aligned} \quad (4)$$

The matrix of this system of linear equations with unknowns $\gamma_{i,j}$, $i = 1, \dots, r$, $j = 0, \dots, s_i - 1$ is:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & 0 & \cdots \\ q_1 & q_1 & q_1 & q_1 & \cdots & q_2 & q_2 & q_2 & q_2 & \cdots \\ q_1^2 & 2q_1^2 & 2^2 q_1^2 & 2^3 q_1^2 & \cdots & q_2^2 & 2q_2^2 & 2^2 q_2^2 & 2^3 q_2^2 & \cdots \\ q_1^3 & 3q_1^3 & 3^2 q_1^3 & 3^3 q_1^3 & \cdots & q_2^3 & 3q_2^3 & 3^2 q_2^3 & 3^3 q_2^3 & \cdots \\ q_1^4 & 4q_1^4 & 4^2 q_1^4 & 4^3 q_1^4 & \cdots & q_2^4 & 4q_2^4 & 4^2 q_2^4 & 4^3 q_2^4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (5)$$

This is a square $k \times k$ matrix. We shall show that it is non-singular by proving that its rows $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_k$ are linearly independent. Assume that there are coefficients $\alpha_1, \dots, \alpha_k$, not all equal to zero, so that $\alpha_1 \mathbf{r}_1 + \alpha_2 \mathbf{r}_2 + \dots + \alpha_k \mathbf{r}_k = \mathbf{0}$. The components of this vector correspond to respective columns in the matrix in the following form:

$$\begin{aligned} \alpha_1 + \alpha_2 q_1 + \alpha_3 q_1^2 + \dots + \alpha_k q_1^{k-1} &= 0 \\ \alpha_1 0 + \alpha_2 q_1 + \alpha_3 2q_1^2 + \dots + \alpha_k (k-1)q_1^{k-1} &= 0 \\ \alpha_1 0 + \alpha_2 q_1 + \alpha_3 2^2 q_1^2 + \dots + \alpha_k (k-1)^2 q_1^{k-1} &= 0 \\ &\vdots \\ \alpha_1 + \alpha_2 q_r + \alpha_3 q_r^2 + \dots + \alpha_k q_r^{k-1} &= 0 \\ \alpha_1 0 + \alpha_2 q_r + \alpha_3 2q_r^2 + \dots + \alpha_k (k-1)q_r^{k-1} &= 0 \\ \alpha_1 0 + \alpha_2 q_r + \alpha_3 2^2 q_r^2 + \dots + \alpha_k (k-1)^2 q_r^{k-1} &= 0 \\ &\vdots \end{aligned}$$

Let $h(x) = \alpha_1 + \alpha_2 x + \alpha_3 x^2 + \dots + \alpha_k x^{k-1}$. Then we see that q_1 is a root of $h(x)$, as well as a root of $h_2(x) = xh_1'(x)$, $h_3(x) = xh_2'(x)$, \dots , h_{s_1} . By Lemma 3.21 $h(x)$ is divisible by $(x - q_1)^{s_1}$. Similarly, $h(x)$ is divisible by $(x - q_2)^{s_2}, \dots, (x - q_r)^{s_r}$. Thus $h(x)$ is divisible by $\prod_{i=1}^r (x - q_i)^{s_i}$. Thus the degree of $h(x)$ is at least $k = \sum_{i=1}^r s_i$, a contradiction to the definition of $h(x)$. Thus there are no such coefficients α_i 's and thus our matrix is non-singular. Therefore the system (4) has a solution for any given right-hand side $(a_0, a_1, \dots, a_{k-1})$. \square

Theorem 3.23 gives us a tool to solve any linear homogeneous recursion with constant coefficients. The main steps are the following.

1. Derive the characteristic polynomial $p(x)$.
2. Compute the roots q_1, \dots, q_r of $p(x)$ with respective multiplicities s_1, \dots, s_r .
3. Solve the system of linear equations (4) (for example by Gaussian Elimination).
4. Write down the explicit solution $a_n = \sum_{i=1}^r \sum_{j=0}^{s_i-1} \gamma_{i,j} n^j q_i^n$.

Let us illustrate this approach with a few examples.

Example 3.24. We solve the recursion

$$a_n = -5a_{n-1} + 6a_{n-2}, \quad a_0 = 3, a_1 = 10.$$

The characteristic polynomial is $x^2 + 5x - 6$, its roots are $q_1 = 1$ and $q_2 = -6$, both with multiplicity 1, i.e., $s_1 = s_2 = 1$. Thus the general solution is $\gamma_{1,0}1^n + \gamma_{2,0}(-6)^n$. Plugging initial values, we obtain

$$\begin{aligned} a_0 &= \gamma_{1,0} + \gamma_{2,0} = 3, \\ a_1 &= \gamma_{1,0} - 6\gamma_{2,0} = 10. \end{aligned}$$

Thus $\gamma_{2,0} = -1$, $\gamma_{1,0} = 4$, and the recursion is solved by $(a_n)_{n \in \mathbb{N}}$ with

$$a_n = 4 - (-6)^n.$$

Example 3.25. We solve the recursion

$$a_n = -4a_{n-1} - 4a_{n-2}, \quad a_0 = 3, a_1 = 10.$$

The characteristic polynomial is $x^2 + 4x + 4$, its roots are $q_1 = -2$, with multiplicity 2, i.e., $s_1 = 2$. Thus the general solution is $\gamma_{1,0}(-2)^n + \gamma_{1,1}n(-2)^n$. Plugging initial values, we obtain

$$\begin{aligned} a_0 &= \gamma_{1,0} = 3, \\ a_1 &= \gamma_{1,0}(-2) + \gamma_{1,1}(-2) = 10. \end{aligned}$$

Thus $\gamma_{1,0} = 3$, $\gamma_{1,1} = -8$, and the recursion is solved by $(a_n)_{n \in \mathbb{N}}$ with

$$a_n = 3(-2)^n - 8n(-2)^n.$$

Example 3.26. We solve the recursion

$$a_n = -4a_{n-2}, \quad a_0 = 6, a_1 = 20.$$

The characteristic polynomial is $x^2 + 4$, its roots are $q_1 = 2i$ and $q_2 = -2i$, both with multiplicity 1, i.e., $s_1 = s_2 = 1$. Thus the general solution is $\gamma_{1,0}(2i)^n + \gamma_{2,0}(-2i)^n$. Plugging initial values, we obtain

$$\begin{aligned} a_0 &= \gamma_{1,0} + \gamma_{2,0} = 6, \\ a_1 &= \gamma_{1,0}(2i) + \gamma_{2,0}(-2i) = 20. \end{aligned}$$

Thus $\gamma_{1,0} = 3 - 5i$, $\gamma_{2,0} = 3 + 5i$. Therefore $a_n = (3 - 5i)(2i)^n + (3 + 5i)(-2i)^n = (2i)^n(3 - 5i + (-1)^n 3 + (-1)^n 5i)$. So, if n is odd, $a_n = (2i)^n(-2 \cdot 5i)$, if n is even $a_n = (2i)^n 6$. Note that $i^n = e^{n \ln i} = e^{\pi n i / 2} = \cos(\pi n / 2) + i \sin(\pi n / 2)$, so for $n = 4k$, $i^n = 1$, $n = 4k + 2$, $i^n = -1$, for $n = 4k + 1$, $i^n = i$, for $n = 4k + 3$, $i^n = -i$.

Hence the recursion is solved by $(a_n)_{n \in \mathbb{N}}$ with

$$a_n = \begin{cases} 5 \cdot 2^{n+1}, & \text{if } n = 4k + 1 \\ 6 \cdot 2^n, & \text{if } n = 4k + 2 \\ -5 \cdot 2^{n+1}, & \text{if } n = 4k + 3 \\ -6 \cdot 2^n, & \text{if } n = 4k + 4. \end{cases}$$

Recall from Example 3.17 that the Fibonacci sequence $(F_n)_{n \in \mathbb{N}}$ solves the recursion

$$F_n = F_{n-1} + F_{n-2}, \quad F_0 = 1, F_1 = 1.$$

We shall now derive an explicit formula for F_n .

Theorem 3.27 (Binet's Formula). *For the Fibonacci numbers we have:*

$$F_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1 + \sqrt{5}}{2} \right)^n - \left(\frac{1 - \sqrt{5}}{2} \right)^n \right)$$

where $\Phi := \frac{1 + \sqrt{5}}{2}$ is called the Golden Ratio.

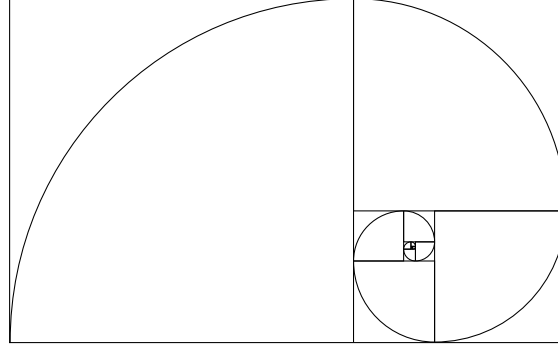


Figure 22: In a drawing like this one, obtained by squares “spiraling” out of the center, the ratio of width and height will approach the golden ratio. Of course, you already know all about spirals, since you watched all [Vi Hart Videos](#), right?

Proof. The recurrence relation is $F_{n+2} = F_{n+1} + F_n$ with starting values $F_0 = 0, F_1 = 1$. So the characteristic polynomial is $p(x) = x^2 - x - 1$. The roots of $p(x)$ are $q_{1,2} = \frac{1 \pm \sqrt{5}}{2}$, i.e.

$$p(x) = \left(x - \frac{1 + \sqrt{5}}{2}\right) \left(x - \frac{1 - \sqrt{5}}{2}\right).$$

The general solution for F_n is therefore (by Theorem 3.23)

$$F_n = \gamma_{1,0} \left(\frac{1 + \sqrt{5}}{2}\right)^n + \gamma_{2,0} \left(\frac{1 - \sqrt{5}}{2}\right)^n.$$

We need to find $\gamma_{1,0}, \gamma_{2,0}$ from initial values³.

$$\begin{aligned} 0 = F_0 &= \gamma_{1,0} \left(\frac{1 + \sqrt{5}}{2}\right)^0 + \gamma_{2,0} \left(\frac{1 - \sqrt{5}}{2}\right)^0 = \gamma_{1,0} + \gamma_{2,0}. \\ 1 = F_1 &= \gamma_{1,0} \left(\frac{1 + \sqrt{5}}{2}\right)^1 + \gamma_{2,0} \left(\frac{1 - \sqrt{5}}{2}\right)^1 \end{aligned}$$

The unique solution is $\gamma_{1,0} = \frac{1}{\sqrt{5}}, \gamma_{2,0} = -\frac{1}{\sqrt{5}}$. □

In the explicit formula we have just shown, the term $\left(\frac{1 - \sqrt{5}}{2}\right)^n \approx (-0.62)^n$ quickly goes to zero. This allows us to ignore it when computing Fibonacci numbers:

Corollary 3.28. F_n is the integer closest to $\frac{1}{\sqrt{5}}\Phi^n$, i.e. $F_n = \left\lfloor \frac{\Phi^n}{\sqrt{5}} + \frac{1}{2} \right\rfloor$.

Proof. It suffices to show that $\frac{1}{\sqrt{5}} \left| \frac{1 - \sqrt{5}}{2} \right|^n < \frac{1}{2}$.

³Any two starting values will do, we could determine $\gamma_{1,0}, \gamma_{2,0}$ from $F_4 = 3, F_5 = 5$ if we wanted to. But using F_0 and F_1 yields the simplest system of equations.

$$\frac{1}{\sqrt{5}} \left| \frac{1 - \sqrt{5}}{2} \right|^n = \frac{|1 - \Phi|^n}{\sqrt{5}} \leq \frac{|1 - \Phi|}{\sqrt{5}} = \frac{1}{\Phi\sqrt{5}} < \frac{2}{\sqrt{5}\sqrt{5}} = \frac{2}{5} < \frac{1}{2}. \quad \square$$

Example 3.29 (Ternary n -string without 20). In Example 3.18 we found starting values and the recurrence but we have yet to determine the values. Recall:

$$t_1 = 3, \quad t_2 = 8, \quad t_{n+2} = 3t_{n+1} - t_n.$$

So $p(x) = (x^2 - 3x + 1)$. The characteristic polynomial has the roots $q_{1,2} = \frac{3 \pm \sqrt{5}}{2}$. So by Theorem 3.23 the general solution is

$$t(n) = \gamma_{1,0} \left(\frac{3 + \sqrt{5}}{2} \right)^n + \gamma_{2,0} \left(\frac{3 - \sqrt{5}}{2} \right)^n.$$

Again we use the initial values to determine $\gamma_{1,0}$ and $\gamma_{2,0}$. For simplicity, use $t_0 = 1$ instead of $t_2 = 8$:

$$\begin{aligned} 1 &= t(0) = \gamma_{1,0} + \gamma_{2,0} \\ 3 &= t(1) = \gamma_{1,0} \cdot \frac{3 + \sqrt{5}}{2} + \gamma_{2,0} \frac{3 - \sqrt{5}}{2} \end{aligned}$$

Solving yields $\gamma_{1,0} = \frac{5+3\sqrt{5}}{10}$, $\gamma_{2,0} = \frac{5-3\sqrt{5}}{10}$ and therefore:

$$t_n = \left(\frac{5+3\sqrt{5}}{10} \right) \left(\frac{3+\sqrt{5}}{2} \right)^n - \left(\frac{5-3\sqrt{5}}{10} \right) \left(\frac{3-\sqrt{5}}{2} \right)^n.$$

3.3.2 Advancement Operator

Remark. This section contains an alternative treatment of recurrence relations in terms of the so-called advancement operator. We present this material, even though it is repetitive and everything is already covered in Section 3.3.1 with the approach of the characteristic polynomial.

We can consider any sequence $(f_n)_{n \in \mathbb{N}}$ as a function $f : \mathbb{N} \rightarrow \mathbb{R}$ where $f(n) = f_n$.

If f fulfills a linear recurrence relation with constant coefficients, then there is a unique way to extend f to \mathbb{Z} such that the recurrence relation is still fulfilled. Take for instance the Fibonacci sequence: Given the starting values there is a unique way to calculate backwards using $F(n-2) = F(n) - F(n-1)$:

n	...	-5	-4	-3	-2	-1	0	1	2	3	4	5	...
F(n)	...	5	-3	2	-1	1	0	1	1	2	3	5	...

This makes f an element of the vector space of all functions from \mathbb{Z} to \mathbb{R} . We define the *advancement operator* A acting on this vector space. It maps a function f to the function Af with:

$$(Af)(n) = f(n+1).$$

An *advancement operator polynomial* is a polynomial in A , for example $3A^2 - A + 6$. This is also an operator which maps f to the function $(3A^2 - A + 6)f$ with:

$$((3A^2 - A + 6)f)(n) = 3f(n+2) - f(n+1) + 6f(n).$$

This notation allows us to write linear recurrence equations with constant coefficients as:

$$p(A)f = g$$

where p is some polynomial. For the Fibonacci numbers we would write $(A^2 - A - 1)F = 0$.

In the following we try to solve equations of this kind. We start with the easiest case of a homogeneous equation and an advancement operator polynomial of degree one.

Lemma 3.30. *For a real number $r \neq 0$ the solution to $(A - r)f = 0$ with initial value $f(0) = c$ is given by $f(n) = cr^n$.*

Proof.

$$\begin{aligned} (A - r)f(n) = 0 &\Leftrightarrow f(n+1) - rf(n) = 0 \\ &\Leftrightarrow f(n+1) = rf(n) \\ &\Leftrightarrow f(n) = r^n \cdot f(0) = r^n \cdot c \end{aligned} \quad \square$$

Raising the difficulty a bit, now consider the advancement operator polynomial $p(A) = A^2 + A - 6 = (A + 3) \cdot (A - 2)$ and the corresponding homogeneous recurrence $p(A)f = (A + 3)(A - 2)f = 0$. Note that solutions of $(A - 2)f = 0$ or $(A + 3)f = 0$ are also solutions of $(A + 3)(A - 2)f = 0$.⁴

Actually, all solutions are sums of such solutions, which means that, using the last Lemma, we know that f is of the form

$$f(n) = c_1(-3)^n + c_22^n.$$

Here, c_1 and c_2 are constants that can be derived from two initial values for f .

It is no coincidence that we found a two dimensional space of functions:

Lemma 3.31. *If $p(A)$ has degree k and its constant term is non-zero, then the set of all solutions f to $p(A)f = 0$ is a k -dimensional subspace of functions from $\mathbb{Z} \rightarrow \mathbb{C}$, parametrized by c_1, \dots, c_k which can be determined by k initial values for f .*

We do not give a formal proof. However, you should not be surprised by the statement. It is clear that set of all solutions form a subspace: If f, g are solutions and $\alpha, \beta \in \mathbb{C} \setminus \{0\}$ then $\alpha f + \beta g$ is also a solution since $p(A)(\alpha f + \beta g) = \alpha p(A)f + \beta p(A)g = \alpha \cdot 0 + \beta \cdot 0 = 0$. It is also intuitive that there are k degrees of freedom: For every choice for values of f on $\{1, 2, \dots, k\}$ the recurrence gives a unique way to extend f (upwards and downwards) to other values.

Generalizing our observation for the advancement operator polynomial $(A + 3)(A - 2)$, we state (also without proof):

⁴There is some low-level notational magic involved here: We need $((A + 3) \cdot (A - 2))f = (A + 3)((A - 2)f) = (A - 2)((A + 3)f)$ where \cdot is multiplication of polynomials and “nothing” is the application of an operator to a function.

Proposition 3.32. *If $p(A) = (A - r_1) \cdot (A - r_2) \cdot \dots \cdot (A - r_k)$ for distinct r_1, \dots, r_k , then all solutions of $p(A)f = 0$ are of the form*

$$f(n) = c_1 r_1^n + c_2 r_2^n + \dots + c_k r_k^n.$$

Every polynomial of degree k has k complex roots, but those roots are not necessarily distinct. The following Theorem handles the case of roots with multiplicity at least two and is therefore the last piece in the puzzle.

From now on, instead of “all solutions are of the form...” we say the “*general solution* is ...”.

Theorem 3.33. *If $r \neq 0$ and $p(A) = (A - r)^k$, then the general solution of the homogeneous system $p(A)f = 0$ is*

$$f(n) = c_1 r^n + n \cdot c_2 r^n + n^2 c_3 r^n + \dots + n^{k-1} c_k \cdot r^n.$$

If $p(A) = q_1(A) \cdot q_2(A)$ and $q_1(A), q_2(A)$ have no root in common, then the general solution of $p(A)f = 0$ is the sum of the general solutions of $q_1(A)f = 0$ and $q_2(A)f = 0$.

Example. Let $p(A) = (A - 1)^5 \cdot (A + 1)^2 \cdot (A - 3)$. Then the general solution to $p(A)f = 0$ is

$$f(n) = c_1 + n \cdot c_2 + n^2 \cdot c_3 + n^3 \cdot c_4 + n^4 \cdot c_5 + c_6(-1)^n + n \cdot c_7(-1)^n + c_8 3^n.$$

Note that the condition that r (the constant term of $p(A)$) may not be zero is no real restriction: An advancement operator polynomial of the form $p(A) \cdot A$ describes the same recurrence as $p(A)$, just at a different index.

Theorem 3.34 (Binet’s Formula). *For the Fibonacci numbers we have:*

$$F_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1 + \sqrt{5}}{2} \right)^n - \left(\frac{1 - \sqrt{5}}{2} \right)^n \right)$$

where $\Phi := \frac{1 + \sqrt{5}}{2}$ is called the *Golden Ratio*.

Proof. The recurrence relation is $F_{n+2} = F_{n+1} + F_n$ with starting values $F_0 = 0, F_1 = 1$. Rewrite this in terms of advancement operator polynomial:

$$f(n + 2) - f(n + 1) - f(n) = 0 \iff \underbrace{(A^2 - A - 1)}_{p(A)} f(n) = 0.$$

The roots of $p(A)$ are $\frac{1 \pm \sqrt{5}}{2}$, i.e.

$$p(A) = \left(A - \frac{1 + \sqrt{5}}{2} \right) \left(A - \frac{1 - \sqrt{5}}{2} \right).$$

The general solution for $f(n)$ is therefore (by Proposition 3.32)

$$f(n) = c_1 \left(\frac{1 + \sqrt{5}}{2} \right)^n + c_2 \left(\frac{1 - \sqrt{5}}{2} \right)^n.$$

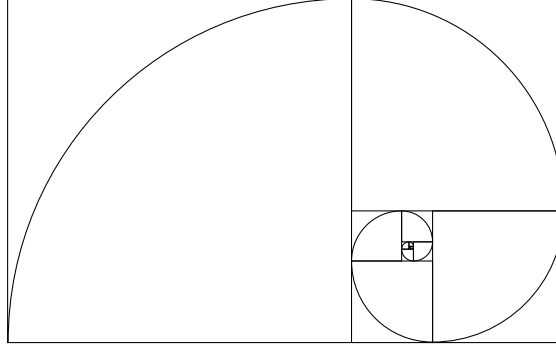


Figure 23: In a drawing like this one, obtained by squares “spiraling” out of the center, the ratio of width and height will approach the golden ratio. Of course, you already know all about spirals, since you watched all [Vi Hart Videos](#), right?

We need to find c_1, c_2 from initial values⁵.

$$\begin{aligned} 0 = f(0) &= c_1 \left(\frac{1 + \sqrt{5}}{2} \right)^0 + c_2 \left(\frac{1 - \sqrt{5}}{2} \right)^0 = c_1 + c_2. \\ 1 = f(1) &= c_1 \left(\frac{1 + \sqrt{5}}{2} \right)^1 + c_2 \left(\frac{1 - \sqrt{5}}{2} \right)^1 \end{aligned}$$

The unique solution is $c_1 = \frac{1}{\sqrt{5}}, c_2 = -\frac{1}{\sqrt{5}}$. □

In the explicit formula we have just shown, the term $(\frac{1-\sqrt{5}}{2})^n \approx (-0.62)^n$ quickly goes to zero. This allows us to ignore it when computing Fibonacci numbers:

Corollary 3.35. F_n is the integer closest to $\frac{1}{\sqrt{5}}\Phi^n$, i.e. $F_n = \left\lfloor \frac{\Phi^n}{\sqrt{5}} + \frac{1}{2} \right\rfloor$.

Proof. It suffices to show that $\frac{1}{\sqrt{5}} \left| \frac{1-\sqrt{5}}{2} \right|^n < \frac{1}{2}$.

$$\frac{1}{\sqrt{5}} \left| \frac{1-\sqrt{5}}{2} \right|^n = \frac{|1-\Phi|^n}{\sqrt{5}} \leq \frac{|1-\Phi|}{\sqrt{5}} = \frac{1}{\Phi\sqrt{5}} < \frac{2}{\sqrt{5}\sqrt{5}} = \frac{2}{5} < \frac{1}{2}. \quad \square$$

Example 3.36 (Ternary n -string without 20). In Example 3.18 we found starting values and the recurrence but we have yet to determine the values. Recall:

$$t_1 = 3, \quad t_2 = 8, \quad t_{n+2} = 3t_{n+1} - t_n.$$

So $p(A)t = (A^2 - 3A + 1)t = 0$. The advancement operator polynomial has the roots $\frac{3 \pm \sqrt{5}}{2}$. So by Proposition 3.32 the general solution is

$$t(n) = c_1 \left(\frac{3 + \sqrt{5}}{2} \right)^n + c_2 \left(\frac{3 - \sqrt{5}}{2} \right)^n.$$

⁵Any two starting values will do, we could determine c_1, c_2 from $F_4 = 3, F_5 = 5$ if we wanted to. But using F_0 and F_1 yields the simplest system of equations.

Again we use the initial values to determine c_1 and c_2 . For simplicity, use $t_0 = 1$ instead of $t_2 = 8$:

$$\begin{aligned} 1 &= t(0) = c_1 + c_2 \\ 3 &= t(1) = c_1 \cdot \frac{3 + \sqrt{5}}{2} + c_2 \frac{3 - \sqrt{5}}{2} \end{aligned}$$

Solving yields $c_1 = \frac{5+3\sqrt{5}}{10}$, $c_2 = \frac{5-3\sqrt{5}}{10}$ and therefore:

$$t_n = \left(\frac{5+3\sqrt{5}}{10} \right) \left(\frac{3+\sqrt{5}}{2} \right)^n - \left(\frac{5-3\sqrt{5}}{10} \right) \left(\frac{3-\sqrt{5}}{2} \right)^n.$$

3.3.3 Non-homogeneous Recurrences

So far we ignored non-homogeneous linear recurrence relations, i.e. those of the form

$$a_n = c_1(n)a_{n-1} + \cdots + c_k(n)a_{n-k} + g(n)$$

for some non-trivial function $g : \mathbb{N} \rightarrow \mathbb{R}$.

Lemma 3.37. *Any two solutions $(a_n)_{n \in \mathbb{N}}$, $(a'_n)_{n \in \mathbb{N}}$ of a non-homogeneous linear recurrence $a_n = f(a_{n-1}, \dots, a_{n-k}) + g(n)$ “differ by” a solution of the corresponding homogeneous recurrence $a_n = f(a_{n-1}, \dots, a_{n-k})$, i.e. there is $(a_n^*)_{n \in \mathbb{N}}$ such that $a_n = a_n^* + a'_n$ for all $n \in \mathbb{N}$ with $a_n^* = f(a_{n-1}^*, \dots, a_{n-k}^*)$ for all $n \geq k$.*

Proof. Just set $a_n^* = a_n - a'_n$ for all $n \in \mathbb{N}$. Then for all $n \geq k$ we have

$$\begin{aligned} f(a_{n-1}^*, \dots, a_{n-k}^*) &= \sum_{i=1}^k c_i(n) a_{n-i}^* = \sum_{i=1}^k c_i(n) (a_{n-i} - a'_{n-i}) \\ &= \sum_{i=1}^k c_i(n) a_{n-i} - \sum_{i=1}^k c_i(n) a'_{n-i} = (a_n - g(n)) - (a'_n - g(n)) \\ &= a_n - a'_n = a_n^*, \end{aligned}$$

as desired. □

This means that to find all solutions of $a_n = f(a_{n-1}, \dots, a_{n-k}) + g(n)$ it suffices to find a single such solution $(a'_n)_{n \in \mathbb{N}}$, the *particular solution*, and all solutions $(a_n^*)_{n \in \mathbb{N}}$ of $a_n = f(a_{n-1}, \dots, a_{n-k})$, the *general solution*.

Then the general solution of the non-homogeneous recurrence is given by $a_n = a_n^* + a'_n$ for $n \in \mathbb{N}$.

Since there is no framework that always works to find particular solutions, this is actually the difficult part.

Example 3.38. Recall from Example 1.6 (and the corresponding problem on the exercise sheet) that when cutting a two dimensional cake into the maximum number of pieces, the n -th cut yields n additional pieces, i.e.

$$s_n = s_{n-1} + n, \quad s_0 = 1 \quad \text{and therefore} \quad s_n = c_1(n)s_{n-1} + g(n)$$

for $c_1(n) = 1$ and $g(n) = n$ for $n \in \mathbb{N}$.

1. Find general solution for the homogeneous recurrence:

$$\begin{aligned} s_n = s_{n-1} &\Rightarrow p(x) = x - 1 \Rightarrow q_1 = 1, s_1 = 1 \\ &\Rightarrow a_n^* = \gamma_{1,0} q_1^n = \gamma_{1,0}. \end{aligned}$$

2. Find particular solution for non-homogeneous recurrence: $s_n = s_{n-1} + n$. We guess that the solution is “similar” to the right hand side. Since the right hand side (i.e. $g(n) = n$) is a polynomial of degree 1 over n , we guess that the solution could be a polynomial of degree 2 over n meaning

$$a'_n = d_1 n^2 + d_2 n + d_3$$

where d_1, d_2, d_3 are to be determined. We plug this guess into the desired recurrence:

$$\begin{aligned} d_1 n^2 + d_2 n + d_3 &= a'_n \stackrel{!}{=} a'_{n-1} + n = d_1 (n-1)^2 + d_2 (n-1) + d_3 + n \\ &= d_1 n^2 + d_2 n + d_3 + (-2d_1 n + d_1 - d_2 + n) \end{aligned}$$

which has the solution $d_1 = d_2 = \frac{1}{2}$ (note that we “lost” d_3 on the way, so it is not needed and we just set it to zero). So we found one particular solution:

$$a'_n = \frac{1}{2} n^2 + \frac{1}{2} n = \frac{n(n+1)}{2} = \binom{n+1}{2}.$$

3. The general solution for the non-homogeneous recurrence is the sum

$$s_n = a_n^* + a'_n = \gamma_{1,0} + \binom{n+1}{2}$$

We still need to determine $\gamma_{1,0}$ using the initial value:

$$\begin{aligned} 1 = s_0 &= \gamma_{1,0} + 0 \Rightarrow \gamma_{1,0} = 1 \\ &\Rightarrow s_n = \binom{n+1}{2} + 1 = \binom{n}{2} + \binom{n}{1} + \binom{n}{0}. \end{aligned}$$

Example 3.39. We solve the recurrence:

$$a_n = 4a_{n-1} - 4a_{n-2} + 3^n + 2n$$

1. General solution for homogeneous recurrence:

$$\begin{aligned} a_n = 4a_{n-1} - 4a_{n-2} &\Rightarrow p(x) = x^2 - 4x + 4 = (x-2)^2 \Rightarrow q_1 = 2, s_1 = 2 \\ &\Rightarrow a_n^* = \gamma_{1,0} q_1^n + \gamma_{1,1} n q_1^n = \gamma_{1,0} 2^n + \gamma_{1,1} n 2^n. \end{aligned}$$

2. Particular solution of the non-homogeneous recurrence:

$$a_n = 4a_{n-1} - 4a_{n-2} + 3^n + 2n$$

Guess something similar to the right hand side. Our attempt is⁶:

$$a'_n = d_1 \cdot 3^n + d_2 n + d_3.$$

⁶We could have guessed another summand of $d_4 \cdot n^2$, but as it turns out, it is not needed.

Which means we need to find d_1, d_2, d_3 such that:

$$\begin{aligned}
d_1 \cdot 3^n + d_2 n + d_3 &= a'_n \stackrel{!}{=} 4a_{n-1} - 4a_{n-2} + 3^n + 2n \\
&= 4d_1 \cdot 3^{n-1} + 4d_2(n-1) + 4d_3 \\
&\quad - 4d_1 \cdot 3^{n-2} - 4d_2(n-2) - 4d_3 + 3^n + 2n \\
&= 8d_1 \cdot 3^{n-2} + 4d_2 + 3^n + 2n \\
&= d_1 \cdot 3^n + d_2 n + d_3 \\
&\quad + (-d_1 \cdot 3^{n-2} - d_2 n - d_3 + 4d_2 + 3^n + 2n) \\
&= d_1 \cdot 3^n + d_2 n + d_3 \\
&\quad + ((9-d_1)3^{n-2} + (2-d_2)n - d_3 + 4d_2)
\end{aligned}$$

and we get $d_1 = 9, d_2 = 2, d_3 = 8$, i.e.

$$a'_n = 3^{n+2} + 2n + 8.$$

3. The general solution of non-homogeneous recurrence is therefore

$$a_n = a_n^* + a'_n = \gamma_{1,0}2^n + \gamma_{1,1}n2^n + 3^{n+2} + 2n + 8.$$

Starting values would now allow us to determine $\gamma_{1,0}$ and $\gamma_{1,1}$ but we omit this here.

3.3.4 Solving Recurrences using Generating Functions

We now examine one example of a recurrence relation with non-constant coefficients.

Example 3.40 (Derangements yet again). Recall from Theorem 1.24 that for $d_n = |D_n|$, the number of derangements on $[n]$, we have

$$d_0 = 1, \quad d_1 = 0, \quad d_2 = 1, \quad d_n = (n-1)(d_{n-1} + d_{n-2}) \quad (\text{for } n \geq 2).$$

This constitutes a homogeneous linear recurrence relation with $k = 2$ and coefficients $c_1(n) = c_2(n) = n-1, g \equiv 0$.

Let $D(x)$ be the exponential generating function for $(d_n)_{n \in \mathbb{N}}$, i.e.

$$D(x) = \sum_{n=0}^{\infty} d_n \frac{x^n}{n!}$$

Then using Observation 3.13 and the recurrence we find:

$$\begin{aligned}
D'(x) &= \sum_{n=0}^{\infty} d_{n+1} \frac{x^n}{n!} = \sum_{n=0}^{\infty} n d_n \frac{x^n}{n!} + \sum_{n=1}^{\infty} n d_{n-1} \frac{x^n}{n!} \\
&= x \sum_{n=1}^{\infty} d_n \frac{x^{n-1}}{(n-1)!} + x \sum_{n=1}^{\infty} d_{n-1} \frac{x^{n-1}}{(n-1)!} \\
&= x D'(x) + x D(x).
\end{aligned}$$

This differential equation has (given $d_1 = 0, d_2 = 1$) the unique solution:

$$D(x) = \frac{e^{-x}}{1-x}.$$

Luckily, we already know the exponential generation functions for e^{-x} and for $\frac{1}{1-x}$ so we can write:

$$\begin{aligned} D(x) &= e^{-x} \cdot \frac{1}{1-x} = \left(\sum_{n=0}^{\infty} (-1)^n \frac{x^n}{n!} \right) \cdot \left(\sum_{n=0}^{\infty} n! \frac{x^n}{n!} \right) \\ &\stackrel{3.12}{=} \sum_{n=0}^{\infty} \left(\sum_{k=0}^n \binom{n}{k} k! (-1)^{n-k} \right) \frac{x^n}{n!}. \end{aligned}$$

So we found yet another proof that $d_n = \sum_{k=0}^n \binom{n}{k} k! (-1)^{n-k}$.

Let us give two more examples showing how we can solve recurrences using generating functions.

Example 3.41. Consider the recurrence

$$a_n = -a_{n-1} + 6a_{n-2} \quad (\text{for } n \geq 2), \quad a_0 = 1, \quad a_1 = 3.$$

With the characteristic polynomial $p(x) = x^2 + x - 6$ in mind we find the following equation for the generating function $F(x) = \sum_{n=0}^{\infty} a_n x^n$ of $(a_n)_{n \in \mathbb{N}}$.

$$\begin{aligned} F(x)(1+x-6x^2) &= \sum_{n=0}^{\infty} a_n x^n + \sum_{n=1}^{\infty} a_{n-1} x^n - 6 \sum_{n=2}^{\infty} a_{n-2} x^n \\ &= a_0 x^0 + a_1 x^1 + a_0 x^1 = 1 + 4x. \end{aligned}$$

Thanks to our approach and using the recurrence for $n \geq 2$, all but a finite number of terms canceled out. In the last step we used the initial values.

Using partial fraction decomposition (not handled here) we can simplify the identity and obtain:

$$\begin{aligned} F(x) &= \frac{1+4x}{1+x-6x^2} = \frac{1+4x}{(1-2x)(1+3x)} = \frac{6}{5} \cdot \frac{1}{1-2x} - \frac{1}{5} \cdot \frac{1}{1+3x} \\ &= \frac{6}{5} \sum_{n=0}^{\infty} (2x)^n - \frac{1}{5} \sum_{n=0}^{\infty} (-3x)^n \end{aligned}$$

where in the last step we applied $\frac{1}{1-y} = \sum_{n=0}^{\infty} y^n$ for $y = 2x$ and $y = -3x$. We can now see the coefficients:

$$a_n = \frac{6}{5} \cdot 2^n - \frac{1}{5} \cdot (-3)^n.$$

Example 3.42. Consider the following non-homogeneous recurrence:

$$b_n = b_{n-1} + 2b_{n-2} + 2^n \quad (\text{for } n \geq 2), \quad b_0 = 2, \quad b_1 = 1.$$

Multiplying the equivalent statement $b_n = b_{n-1} + 2b_{n-2} + 2^n$ with x^n for each $n \geq 2$ and adding yields a sum in which we can rewrite the terms:

$$\underbrace{\sum_{n=2}^{\infty} b_n x^n}_{F(x) - b_0 - x b_1} - \underbrace{\sum_{n=2}^{\infty} b_{n-1} x^n}_{x \cdot (F(x) - b_0)} - 2 \underbrace{\sum_{n=2}^{\infty} b_{n-2} x^n}_{x^2 F(x)} = \underbrace{\sum_{n=2}^{\infty} 2^n x^n}_{\frac{1}{1-2x} - 1 - 2x}.$$

Solving for $F(x)$ yields:

$$\begin{aligned} F(x)(1-x-2x^2) &= \frac{1}{1-2x} + 1 - 3x \\ \Rightarrow F(x) &= \frac{2-5x+6x^2}{(1-2x)(1-x-2x^2)} \end{aligned}$$

From here, apply partial fraction decomposition **with a method of your choice**, to get

$$F(x) = -\frac{1}{9} \frac{1}{1-2x} + \frac{2}{3} \frac{1}{(1-2x)^2} + \frac{13}{9} \frac{1}{1+x}.$$

From this we can obtain the coefficients again using identities we know.

$$\begin{aligned} F(x) &= -\frac{1}{9} \sum_{n=0}^{\infty} 2^n x^n + \frac{2}{3} \left(\sum_{n=0}^{\infty} 2^n x^n \right)^2 + \frac{13}{9} \sum_{n=0}^{\infty} (-1)^n x^n. \\ &= -\frac{1}{9} \sum_{n=0}^{\infty} 2^n x^n + \frac{2}{3} \sum_{n=0}^{\infty} (n+1) 2^n x^n + \frac{13}{9} \sum_{n=0}^{\infty} (-1)^n x^n. \end{aligned}$$

The coefficients are therefore:

$$b_n = -\frac{1}{9} 2^n + \frac{2}{3} (n+1) 2^n + \frac{13}{9} (-1)^n.$$

4 Partitions

4.1 Partitioning $[n]$ – the set on n elements

A partition of $[n]$ into given by its parts A_1, \dots, A_k with

$$A_i \neq \emptyset \text{ (for } i = 1, \dots, k), \quad A_i \cap A_j = \emptyset \text{ (for } i \neq j), \quad \bigcup_{i=1}^k A_i = [n].$$

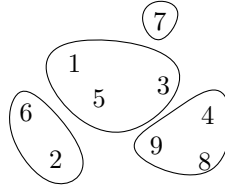


Figure 24: A partition of $[9]$ into 4 sets.

The parts are unlabeled, i.e. if two partitions use the same parts A_i only in a different order, we consider them to be identical. If we fix the number of parts, i.e. if we want to count partitions of $[n]$ into exactly k non-empty parts, then their number is given by $s_k^{II}(n)$ as we have already seen in Section 1.5.2 where we considered arrangements of n labeled balls in k unlabeled boxes with at least one ball per box.

We define the *Bell Number* as

$$B_n = \sum_{k=0}^n s_k^{II}(n).$$

It counts the total number of partitions of $[n]$ (into an arbitrary number of sets). Don't get confused over the special case of $n = 0$: There is exactly one partition of \emptyset into non-empty parts: $\emptyset = \bigcup_{A \in \emptyset} A$. Every $A \in \emptyset$ is non-empty, since no such A exists. So we also have $B_0 = s_0^{II}(0) = 1$.

A different way to define the Bell numbers is to consider a square free number $k \in \mathbb{N}$, meaning

$$k = p_1 \cdot p_2 \cdot \dots \cdot p_n \quad \text{for distinct primes } p_1, \dots, p_n.$$

The number of ways to write k as product of integers bigger than one is exactly B_n . Take for instance $k = 2 \cdot 3 \cdot 5 \cdot 7$, then some ways of writing k would be:

$$k = 6 \cdot 35, \quad k = 2 \cdot 5 \cdot 21, \quad k = 210$$

which directly corresponds to the partitions:

$$\{2, 3, 5, 7\} = \{2, 3\} \cup \{5, 7\}, \quad \{2, 3, 5, 7\} = \{2\} \cup \{5\} \cup \{3, 7\}, \quad \{2, 3, 5, 7\} = \{2, 3, 5, 7\}.$$

In the following we try to find an explicit formula for B_n . We start by finding a recursion:

Theorem 4.1.

$$B_n = \sum_{k=0}^{n-1} \binom{n-1}{k} B_k \quad (n \geq 1).$$

Proof. Every partition of $[n]$ has one part that contains the number n . In addition to n this part contains k other numbers (for some $0 \leq k \leq n-1$). The remaining $n-1-k$ elements are partitioned arbitrarily. From this correspondence we obtain the desired identity:

$$B_n = \sum_{k=0}^{n-1} \binom{n-1}{k} B_{n-1-k} = \sum_{k=0}^{n-1} \binom{n-1}{n-1-k} B_{n-1-k} = \sum_{k=0}^{n-1} \binom{n-1}{k} B_k. \quad \square$$

Theorem 4.2 (Dobinski's Formula).

$$B_n = \frac{1}{e} \sum_{k=0}^{\infty} \frac{k^n}{k!}.$$

Proof. Consider the exponential generating function $B(x)$ of $(B_n)_{n \in \mathbb{N}}$.

$$B(x) = \sum_{n=0}^{\infty} B_n \frac{x^n}{n!}$$

Using the recursion from Theorem 4.1 we see

$$\begin{aligned} B'(x) &= \sum_{n=0}^{\infty} B_{n+1} \frac{x^n}{n!} = \sum_{n=0}^{\infty} \left(\sum_{k=0}^n \binom{n}{k} B_k \right) \frac{x^n}{n!} \\ &\stackrel{3.12}{=} \left(\sum_{n=0}^{\infty} B_n \frac{x^n}{n!} \right) \cdot \left(\sum_{n=0}^{\infty} 1 \frac{x^n}{n!} \right) = B(x) \cdot e^x. \\ \Rightarrow B(x) &= e^{e^x - 1} \quad (\text{using } B_0 = B(0) = 1). \\ &= \frac{1}{e} e^{e^x} = \frac{1}{e} \sum_{k=0}^{\infty} \frac{(e^x)^k}{k!} = \frac{1}{e} \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{n=0}^{\infty} \frac{(xk)^n}{n!} \\ &= \sum_{n=0}^{\infty} \left(\frac{1}{e} \sum_{k=0}^{\infty} \frac{k^n}{k!} \right) \frac{x^n}{n!} \quad \square \end{aligned}$$

4.1.1 Non-Crossing Partitions

Imagine the elements of $[n]$ to be laid out in a circular way.

Two disjoint sets $A, B \subseteq [n]$ are crossing if there are numbers $i < j < k < l \in [n]$ such that $\{i, k\} \subseteq A, \{j, l\} \subseteq B$.

A *non-crossing* partition is a partition in which the parts are pairwise non-crossing. In cyclic drawings the notion is very intuitive. See Figure 25 for examples.

We denote the number of non-crossing partitions of $[n]$ by NC_n . We prove now that NC_n is equal to C_n , the n -th Catalan number. We already came across these numbers in Example 3.5, where we counted well formed parenthesis expressions.

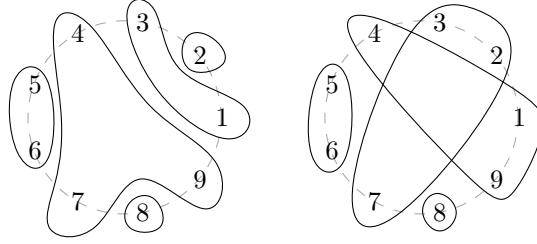


Figure 25: A non-crossing partition of $[9]$ on the left and a crossing partition of $[9]$ on the right.

Theorem 4.3.

$$\text{NC}_n = C_n = \frac{1}{n+1} \binom{2n}{n}.$$

Proof. Recall the values $C_0 = 1, C_1 = 1, C_2 = 2$ (corresponding to the parenthesis expressions “”, “()”, “()()”, “(())”) and the recursion

$$C_{n+1} = \sum_{k=0}^n C_k C_{n-k}.$$

It suffices to prove that the sequence $(\text{NC}_n)_{n \in \mathbb{N}}$ has the same starting values and satisfies the same recursion.

It is easy to check that $\text{NC}_0 = 1, \text{NC}_1 = 1, \text{NC}_2 = 2$. Actually those numbers are just the Bell numbers since partitions of $[n]$ can only be crossing for $n \geq 4$.

We now have to prove the recursion

$$\text{NC}_{n+1} = \sum_{k=0}^n \text{NC}_k \cdot \text{NC}_{n-k}.$$

To this end, consider any non-crossing partition \mathcal{P} of $[n+1]$. The last element $n+1$ is in some part $S \subseteq [n+1]$. Let k be the biggest number in S other than $n+1$ if such an element exists and $k = 0$ if $S = \{n+1\}$. Now observe that in the partition \mathcal{P} , every part contains either only numbers that are bigger than k or only numbers that are at most k : Otherwise, such a part would cross S .

This means that \mathcal{P} decomposes into a non-crossing partition of $[k]$ and a non-crossing partition of $\{k+1, \dots, n\}$. Here we ignored $n+1$: It must be in the same part as k and will never produce a crossing if there has not already been one. Such a decomposition is unique: Every non-crossing partition of $[n+1]$ uniquely decomposes and every pair of non-crossing partitions of $[k]$ and $\{k+1, \dots, n\}$ corresponds to a non-crossing partition of $[n+1]$.

This proves the claimed recursion and therefore the Theorem. \square

4.2 Partitioning n – the natural number

We are interested in ways to write n as the sum of positive natural numbers. We would say, for instance, that

$$n = 17 = 5 + 5 + 4 + 3.$$

is a *partition* of $n = 17$ into the (unlabeled) *parts* 5, 5, 4 and 3.

Alternatively we write $n = \lambda = (5, 5, 4, 3)$ and say λ is the partition of n (even though λ is a sorted tuple) hoping this will not be confusing.

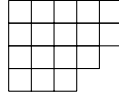
We already considered this in Section 1.5.4 where we counted arrangements of n unlabeled balls in k unlabeled boxes and at least one ball per box. The number of such arrangements is given by $p_k(n)$. We established a recursive formula

$$p_k(n) = \begin{cases} 0 & \text{if } k > n, \\ 0 & \text{if } n \geq 1, k = 0, \\ 1 & \text{if } n = k = 0, \\ p_k(n - k) + p_{k-1}(n - 1) & \text{if } 1 \leq k \leq n. \end{cases}$$

We define the total number of partitions of n (into an arbitrary number of parts) as the *partition function*

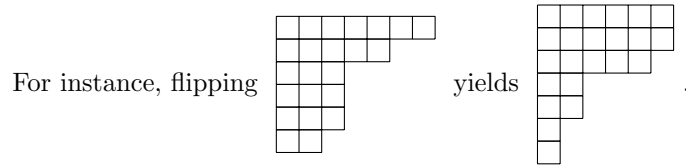
$$p(n) = \sum_{k=0}^n p_k(n).$$

We illustrate a partition by its *Ferrer diagram* (also called *Young diagram*). It consists of rows of squares (left aligned) corresponding to the parts (largest part in the topmost row). Consider for instance the partition $\lambda = (5, 5, 4, 3)$ of $n = 17$ which has the Ferrer diagram



Observation 4.4. The number of partitions of n into at most k parts is equal to the number of partitions of n with parts of size at most k .

Proof. The Ferrer diagrams for partitions with at most k parts are those with at most k rows. The Ferrer diagrams for partitions with parts of size at most k are those with at most k columns. Clearly there is a bijection between those sets of diagrams: Flip them along the diagonal!



Which means the diagram for the partition $\lambda = (7, 5, 3, 3, 3, 2)$ is mapped to the diagram for the partition $\lambda = (6, 6, 5, 2, 2, 1, 1)$. \square

Formal Proof. The bijection on the diagrams corresponds to a bijection on the partitions that maps the partition $n = (\lambda_1, \dots, \lambda_k)$ with biggest part l to the *conjugate partition* $n = (\lambda_1^*, \dots, \lambda_l^*)$ defined as

$$\lambda_i^* = |\{j \mid \lambda_j \geq i\}|. \quad \square$$

Theorem 4.5 (Euler).

$$\sum_{n=0}^{\infty} p(n)x^n = \prod_{k=1}^{\infty} \frac{1}{1-x^k}.$$

Proof. We boldly rewrite the infinite product as an infinite product of infinite sums using the identity: $\frac{1}{1-y} = \sum_{n=0}^{\infty} y^n$.

$$\begin{aligned} \prod_{k=1}^{\infty} \frac{1}{1-x^k} &= \frac{1}{1-x} \cdot \frac{1}{1-x^2} \cdot \frac{1}{1-x^3} \cdot \dots \\ &= \left(\sum_{n_1=0}^{\infty} x^{n_1} \right) \left(\sum_{n_2=0}^{\infty} x^{2n_2} \right) \left(\sum_{n_3=0}^{\infty} x^{3n_3} \right) \dots \end{aligned}$$

In an expansion of this, the coefficient in front of x^n is just the number of ways we can choose n_1, n_2, n_3, \dots such that $\sum_{i \geq 0} i \cdot n_i = n$. But such choices exactly correspond to partitions of n , i.e. the coefficient is $p(n)$.

Take for instance our favorite partition $n = 17 = 5 + 5 + 4 + 3$. It corresponds to choosing $n_5 = 2$, $n_4 = 1$, $n_3 = 1$ and all other indices n_i as 0. This proves the claim. \square

Define $p_{\text{odd}}(n)$ as the number of partitions of n into odd parts and $p_{\text{dist}}(n)$ as the number of partitions of n into distinct parts.

Take for instance $n = 7$. We have $p_{\text{odd}}(7) = 5$ since

$$7 = 1+1+1+1+1+1+1, \quad 7 = 1+1+1+1+3, \quad 7 = 1+3+3, \quad 7 = 1+1+5, \quad 7 = 7,$$

and $p_{\text{dist}}(7) = 5$ since

$$7 = 3 + 4, \quad 7 = 1 + 2 + 4, \quad 7 = 2 + 5, \quad 7 = 1 + 6, \quad 7 = 7.$$

This is no coincidence, in fact, in the problem class you already saw:

$$\sum_{n=0}^{\infty} p_{\text{odd}}(n)x^n = \prod_{k=0}^{\infty} \frac{1}{1-x^{2k+1}} = \prod_{k=0}^{\infty} (1+x^k) = \sum_{n=0}^{\infty} p_{\text{dist}}(n)x^n.$$

This proves $p_{\text{odd}} = p_{\text{dist}}$, but we will prove it yet again, this time by constructing a bijection.

Theorem 4.6. $p_{\text{odd}} = p_{\text{dist}}$.

Proof. Let $n = \lambda_1 + \lambda_2 + \dots + \lambda_k$ be a partition of n into distinct parts. We separate the powers of 2 from the λ_i , i.e. we write:

$$n = u_1 2^{a_1} + u_2 2^{a_2} + \dots + u_k 2^{a_k}$$

for odd numbers u_i and $\lambda_i = u_i 2^{a_i}$. Note that the u_i need no longer be distinct, for example if $\lambda_1 = 5, \lambda_2 = 10$ we would have $u_1 = u_2 = 5$. We sort the summands according to the u_i which gives

$$n = \mu_1 \cdot \underbrace{(2^{\alpha_1} + 2^{\alpha_2} + \dots + 2^{\alpha_{l_1}})}_{r_1} + \mu_2 \cdot \underbrace{(2^{\beta_1} + 2^{\beta_2} + \dots + 2^{\beta_{l_2}})}_{r_2} + \dots \quad (\star)$$

where the values μ_i are all distinct and take the roles of u_i , in particular, the following multisets coincide

$$\{u_1, u_2, \dots, u_k\} = \underbrace{\{\mu_1, \mu_1, \dots, \mu_1\}}_{l_1 \text{ times}} \underbrace{\{\mu_2, \mu_2, \dots, \mu_2\}}_{l_2 \text{ times}} \dots$$

Note that the values r_i are sums of *distinct* powers of 2.

For the bijection, we map the original partition into distinct parts to the following partition into odd parts

$$n = \underbrace{\mu_1 + \mu_1 + \dots + \mu_1}_{r_1 \text{ times}} + \underbrace{\mu_2 + \mu_2 + \dots + \mu_2}_{r_2 \text{ times}} + \dots$$

We still need to argue that this constitutes a bijection. To do so, we explain the inverse mapping.

Given a partition into odd numbers with repetition numbers r_i , there is a unique way to write these r_i as sums of distinct powers of 2, the binary expansion of r_i . This gets us to the situation (\star) and from there we get back to a partition into distinct parts by multiplying out. \square

Example. To illustrate the last Theorem, take the partition into this distinct parts:

$$26 = 12 + 6 + 4 + 3 + 1.$$

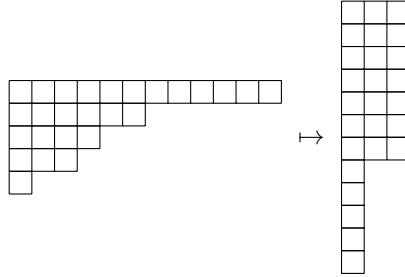
separate the powers of two and sort the terms according to the odd numbers that remain:

$$\begin{aligned} 26 &= 3 \cdot 2^2 + 3 \cdot 2^1 + 1 \cdot 2^2 + 3 \cdot 2^0 + 1 \cdot 2^0 \\ &= 3(2^0 + 2^1 + 2^2) + 1(2^0 + 2^2) \\ &= 3 \cdot 7 + 1 \cdot 5 \end{aligned}$$

which yields the partition

$$26 = 3 + 3 + 3 + 3 + 3 + 3 + 3 + 1 + 1 + 1 + 1 + 1.$$

All steps are reversible, as there is only one way to write 5 and 7 as sums of distinct powers of two. In terms of Ferrer diagrams we have mapped:



Now define $p_d^{\text{even}}(n)$ and $p_d^{\text{odd}}(n)$ as the numbers of partitions of n into an even number of distinct parts and an odd number of distinct parts, respectively. We have, for instance

$$\begin{aligned} p_d^{\text{even}}(7) &= \#\{“1+6”, “2+5”, “3+4”\} = 3, \\ p_d^{\text{odd}}(7) &= \#\{“1+2+4”, “7”\} = 2. \end{aligned}$$

Apparently, these numbers can differ, but we now show that they can differ by at most 1, and characterize when this is the case.

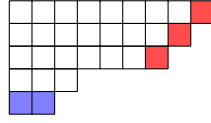
To this end, define for $k \in \mathbb{Z}$ the *pentagonal number* $w_k := \frac{(3k-1)k}{2}$. Note that with this definition $w_{-k} = \frac{(3k+1)k}{2}$. Some values are:

k	\dots	-3	-2	-1	0	1	2	3	\dots
w_k	\dots	15	7	2	0	1	5	12	\dots

Lemma 4.7.

$$p_d^{\text{even}}(n) - p_d^{\text{odd}}(n) = \begin{cases} (-1)^k & \text{if } n = w_k \text{ for some } k \in \mathbb{Z}, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Consider the Ferrer Diagrams for a partition into distinct parts. In it, two rows may never have the same length. Define the *slope* S of such a diagram to be a maximal staircase going diagonally down starting from the top-right square and define the bottom B as the last row of the diagram. The following diagram has a slope of length 3 (highlighted in red) and a bottom of size 2 (highlighted in blue):



Note that, in a few special diagrams, B and S may have a single square in common. We define the set $\Delta := B \cap S$, it contains the single common square if it exists, and is empty otherwise. We will be sloppy in notation using B , S and Δ to simultaneously denote the set and the size of the set.

Now distinguish three types partitions corresponding to three types of diagrams where:

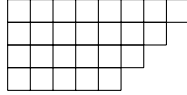
- Type 1: $S \geq B + \Delta$
- Type 2: $S < B - \Delta$
- Type 3: $B - \Delta \leq S < B + \Delta$

Note that the latter case can only occur for $\Delta = 1$ and $S \in \{B, B - 1\}$.

Claim. (i) Type 3 partitions can only occur if $n = w_k$ for some $k \in \mathbb{Z}$.

(ii) Conversely, if $n = w_k$ for some $k \in \mathbb{Z}$, then there is exactly one Type 3 partition of n .

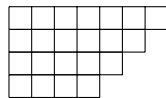
Proof of (i). Consider the case $k := S = B - 1$ first, where the diagram looks like this (for $k = 4$):



which gives

$$n = k^2 + \sum_{i=1}^k i = k^2 + \frac{(k+1)k}{2} = \frac{(3k+1)k}{2} = w_{-k}.$$

The other case is $k := S = B$ meaning the diagram looks like this (for $k = 4$)



which gives

$$n = w_{-k} - k = \frac{(3k+1)k}{2} - \frac{2k}{2} = \frac{(3k-1)k}{2} = w_k.$$

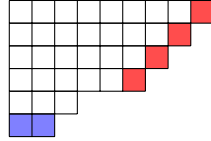
Proof of (ii). The existence is already clear from our proof of (i): We found for arbitrary $k \in \mathbb{N}$ a type 3 partition of w_{-k} and w_k . For the uniqueness, note that no two diagrams of type 3 have the same size: We can iterate through all of them by alternatingly adding a column and a row.

Next we prove the following claim:

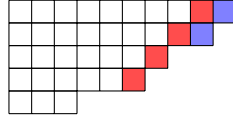
Claim (iii). For every $k \in \mathbb{N}$ there is a bijection between Type 1 partitions on k rows and Type 2 partitions on $k-1$ rows.

Note that from this the Theorem follows, since it guarantees a bijection that maps every partition to a partition with one row more or one row less, so every partition into an even number of parts is mapped to a partition into an odd number of parts and vice versa. This shows that the number of even and odd partitions must coincide. The only disturbance can be the single type 3 partition that exists if $n = w_k$ and is even if k is even and odd if k is odd.

Proof of (iii). Consider a Type 1 partition, its slope is at least as large as its bottom, maybe like this:



We take away the bottom and distribute the squares among the first $|B|$ rows of the diagram, like this:



There is enough room to do this (since the slope was at least as big as the bottom) and the resulting partition is a partition into distinct parts. The size of the bottom has increased and the size of the new slope is the size of the old bottom. Therefore the new diagram is of type 2 or type 3 and, looking more closely, type 3 can actually not occur as result of our operation, so it really is of type 2. The inverse operation is to take the current slope and create a new row from it. After checking that this maps type 2 partitions to type 1 partitions we are done. \square

Theorem 4.8 (Euler's Pentagonal Number Theorem).

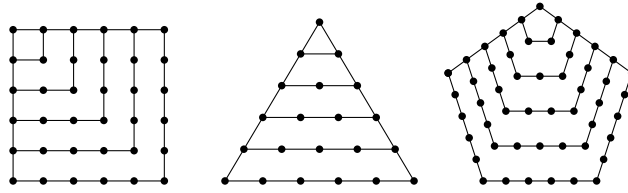
$$\prod_{k=1}^{\infty} (1 - x^k) = \sum_{k=-\infty}^{\infty} (-1)^k x^{w_k} = 1 + \sum_{k=1}^{\infty} (-1)^k (x^{w_k} + x^{w_{-k}}).$$

Proof. Multiply out the left hand side. The coefficient for x^n counts the number of partitions of n into distinct parts, however, partitions into an odd number of parts are counted with negative sign. Therefore

$$\prod_{k=1}^{\infty} (1 - x^k) = \sum_{n=0}^{\infty} (p_d^{\text{even}}(n) - p_d^{\text{odd}}(n)) x^n \stackrel{\text{Thm 4.7}}{=} \sum_{k=-\infty}^{\infty} (-1)^k x^{w_k}.$$

□

We have yet to explain why w_k are called pentagonal numbers. It is basically analogous to square numbers and triangle numbers: All occur naturally when filling a two-dimensional region with dots.



We start with one dot, and say it is “layer 0”, then add layer by layer filling an area of the respective kind. For pentagonal numbers, there are four dots in layer 1, seven dots in layer 2 and so on. Generally, in layer i there are $3i + 1$ dots since there are three sides with $i + 1$ dots each, but two dots are shared by two sides. In a drawing with k layers there are:

$$\sum_{i=0}^{k-1} 3i + 1 = k + 3 \left(\sum_{i=0}^{k-1} i \right) = k + 3 \binom{k}{2} = \frac{(3k-1)k}{2} = w_k.$$

4.3 Young Tableau

A Young tableau T is a Ferrer diagram of a partition λ of n that is filled with all numbers $1, \dots, n$. For instance if $\lambda = (7, 6, 5, 3, 2, 2, 1)$ is a partition of 26:

11	26	23	13	19	14	24
8	18	9	15	22	21	
5	1	2	3	6		
17	4	7				
10	12					
16	25					
20						

We say T is a *tableau of λ* or T *has shape λ* . There are $n!$ tableaux of shape λ .

Definition 4.9. A *standard Young tableau* is a Young tableau where the numbers in the rows are increasing from left-to-right and the numbers in the columns are increasing from top-to-bottom, take for instance the following standard

Young tableau:

1	2	3	6	12	16	25
4	7	8	17	18	21	
5	11	13	19	26		
9	15	22				
10	23					
14	24					
20						

From now on we only consider standard Young tableaux.

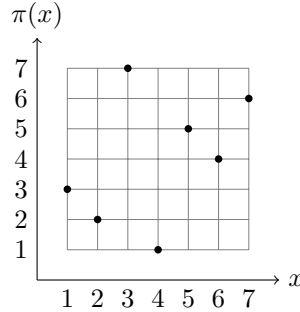
Theorem 4.10 (Robinson-Schensted-Correspondence). *There is a bijection between the permutations of $[n]$ and the set*

$$\bigcup_{\lambda \text{ is partition of } n} \{(T_1, T_2) \mid T_1 \text{ and } T_2 \text{ are standard Young tableaux of shape } \lambda\}.$$

i.e. the set of ordered pairs (T_1, T_2) where T_1 and T_2 are standard Young tableaux of the same shape.

Proof. We will see a geometric construction that builds, given a permutation π , a corresponding pair of standard Young tableaux.

Let π be a permutation of $[n]$ and $X(\pi) = \{(i, \pi(i)) \mid i = 1, \dots, n\}$ the corresponding point set in the plane. For instance, for the permutation $\pi = 3271546$ of $[7]$ the point set $X(\pi)$ is:

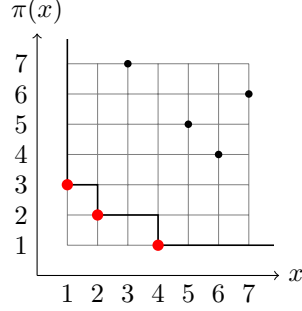


A point p is *minimal* if there is no other point that is both to the left and below of p . The set of all minimal points in X is therefore

$$\min(X) = \{(x, y) \in X \mid \forall (x', y') \in X \setminus \{x, y\}: x' > x \text{ or } y' > y\}.$$

The *shadowline* $S(X)$ for a point set X is the weakly decreasing rectilinear line through all points in $\min(X)$ and convex bends exactly in $\min(X)$. Here, by convex we mean \sqcup and by concave we mean \sqcap . In the following drawing,

$\min(X)$ (red) and $S(X)$ (black) are shown.



In case you like pseudo-code, a concise algorithmic description of our construction is given below. It does not yet provide the pair of tableaux we want, but all the auxiliary objects we require to define them. A detailed description in words follows.

Algorithm 1 Geometric Variant of Robinson-Schensted-Correspondence

```

 $X_1 \leftarrow X(\pi)$ 
 $i \leftarrow 0$ 
while  $X_{i+1} \neq \emptyset$  do
   $i \leftarrow i + 1$ 
   $X'_i \leftarrow X_i$ 
   $j \leftarrow 0$ 
  while  $X'_i \neq \emptyset$  do
     $j \leftarrow j + 1$ 
     $S_i^j \leftarrow S(X'_i)$ 
     $X'_i \leftarrow X'_i \setminus \min(X'_i)$ 
  end while
   $n_i \leftarrow j$ 
   $X_{i+1} \leftarrow \text{convex bends of } S_i^1, \dots, S_i^{n_i}$ 
end while
 $m \leftarrow i$ 

```

The algorithm proceeds in phases (counted by the variable i), the total number of phases m is not known beforehand (but bounded by n). We start every phase i with a non-empty point set X_i . For X_i we construct a sequence of shadowlines $S_i^1, \dots, S_i^{n_i}$ where the j -th shadowline is taken for the point set consisting of those points from X_i that were not used for previous shadowlines.

$$S_i^1 = S(X_i), \quad S_i^2 = S(X_i \setminus S_i^1). \quad \text{In general: } S_i^j = S(X_i \setminus \bigcup_{k=1}^{j-1} S_i^k).$$

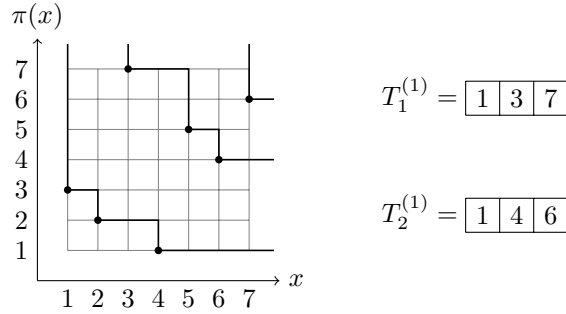
The i -th phase ends as soon as all points from X_i were contained in one of the shadowlines $S_i^1, S_i^2, \dots, S_i^{n_i}$.

The shadowlines of phase i determine the i -th row of the tableaux T_1 and T_2 we want to construct. Let x_i^j be the x -coordinate of first segment of the shadowline S_i^j , i.e. the x -coordinate at which S_i^j leaves the picture on the top and y_i^j the y -coordinate of last segment of S_i^j , i.e. the y -coordinate at which S_i^j leaves the picture on the right. Then the i -th row of T_1 and T_2 consists of the numbers $x_i^1, \dots, x_i^{n_i}$ and $y_i^1, \dots, y_i^{n_i}$, respectively.

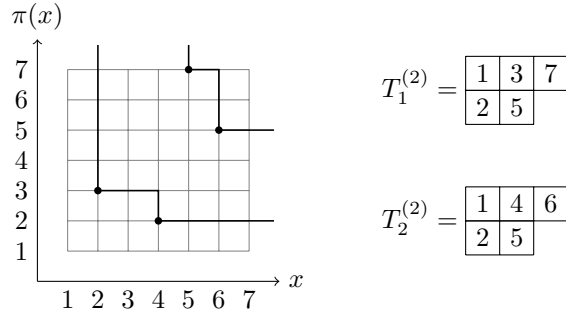
Shadowlines may contain concave bends (they do iff they have two or more convex bends). The set X_{i+1} is defined as the set of all concave bends occurring in the shadowlines $S_i^1, \dots, S_i^{n_i}$. If X_{i+1} is empty, we are done, otherwise, it serves as point set for phase $i+1$. *(proof continues later)*

Before we verify that the construction yields the bijection we desire, we give an example.

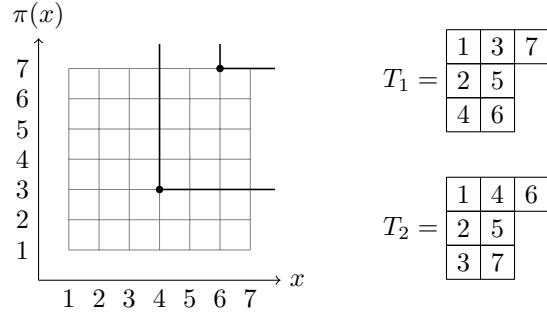
Example. Consider again $\pi = 3271546$. The first phase of the algorithm will find three shadowlines S_1^1, S_1^2, S_1^3 . They leave the diagram on x positions 1, 3 and 7 and on y positions 1, 4 and 6, giving rise to the partial tableaux as shown.



There are four concave bends on these shadowlines which give the point set for the next phase: $X_2 = \{(2, 3), (4, 2), (5, 7), (6, 5)\}$. In phase 2 we get two shadowlines and add corresponding second lines to the tableaux.



There are still convex bends, so we proceed with phase 3



No concave bends were made this time, our construction is done. The pair (T_1, T_2) is the result.

We now establish, in a series of claims, that the construction constitutes a bijection between permutations and pairs of Young tableaux of the same shape as desired. We use the notion of a *chain*, which is a subset Y of a point set X such that Y is *increasing*, i.e. $Y = \{(x_1, y_1), (x_2, y_2), \dots, (x_k, y_k)\}$ with $y_1 < y_2 < \dots < y_k$ and $x_1 < x_2 < \dots < x_k$.

Claim. The number n_i of shadow lines in phase i is the length of the longest chain in X_i .

Proof of Claim: Let Y be a longest chain in X_i . No shadowline can contain more than one point of Y , since shadowlines are decreasing while Y is increasing. This shows $n_i \geq |Y|$.

Now observe that for the element $\mu \in Y$ with smallest x - and y -coordinate we have $\mu \in \min(X_i)$, otherwise there would be a point to the left and below of μ enabling us to extend Y . Therefore, the first shadowline S_i^1 will contain μ (and, indeed, an element of every longest chain) so the longest chain in $X_i \setminus S_i^1$ will be of size $|Y| - 1$. The next shadowline will contain another point of Y and, by induction, we conclude that $n_i \leq |Y|$ since $X_{1+|Y|}$ will only contain chains of size 0, meaning $X_{1+|Y|}$ is empty. $\square(\text{claim})$

Claim. The tableaux have valid shape, i.e. $n_{i+1} \leq n_i$ for all $i = 1, \dots, m-1$.

Proof of Claim: Let Y be a longest chain in X_{i+1} , consisting of some positions of concave bends of the shadowlines from the previous phase i . Since the concave bends of every shadowline are decreasing, Y can only contain one concave bend from each shadowline of phase i . This means $|Y| \leq n_i$. We already know $|Y| = n_{i+1}$ from the previous claim, so we are done. $\square(\text{claim})$

Claim. The rows of T_1 and T_2 are increasing.

Proof of Claim: This is equivalent to saying: Shadowlines constructed later in the phase will leave the picture further to the right and further to the top. This is obvious by construction: Every shadowline goes through the unused point with smallest x coordinate and the unused point with smallest y coordinate, making that point unavailable for later shadowlines. $\square(\text{claim})$

Claim. The columns of T_1 and T_2 are increasing.

Proof of Claim: We consider two consecutive phases i and $i+1$ and show that in the step from row i to row $i+1$ every column is increasing.

In phase i the shadowlines $S_i^1, \dots, S_i^{n_i}$ were constructed with corresponding sets of concave bends $B_i^1, \dots, B_i^{n_i}$ (some of which may be empty) that together form the set X_{i+1} . In the following we implicitly use that shadow lines of the same phase are non-crossing. We have $B_i^1 \subseteq \min(X_{i+1})$, meaning that the points B_i^1 will be “consumed” by the first shadowline S_{i+1}^1 of phase $i+1$. Therefore S_{i+1}^2 will be constructed from a subset of $B_i^2, \dots, B_i^{n_i}$, and will use up all remaining points from B_i^2 (if any remain). Generally, S_{i+1}^j will be constructed from a subset of $B_i^j, \dots, B_i^{n_i}$. The x -coordinate of the leftmost point of S_{i+1}^j is therefore at least the x -coordinate of the leftmost concave bend of S_i^j and therefore to the right of the leftmost point of S_i^j . This proves that columns of T_1 are increasing, for T_2 do the same argument for y coordinates. $\square(\text{claim})$

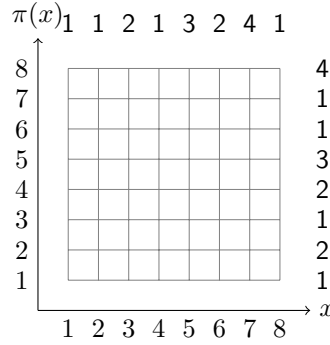
So far we proved that our map is well-defined, i.e. it gives rise to a pair of standard Young tableaux for any permutation $\pi \in S_n$. The final step is to show that the map constitutes a bijection.

Claim. There is an inverse map, i.e. from any pair (T_1, T_2) of standard Young tableaux of the same shape we can recover a corresponding permutation $\pi \in S_n$.

Proof. We demonstrate the inverse procedure with an example, hoping the general case will be apparent from this. Consider the following pair of Young tableaux:

$$T_1 := \begin{array}{|c|c|c|c|} \hline 1 & 2 & 4 & 8 \\ \hline 3 & 6 & & \\ \hline 5 & & & \\ \hline 7 & & & \\ \hline \end{array}, \quad T_2 := \begin{array}{|c|c|c|c|} \hline 1 & 3 & 6 & 7 \\ \hline 2 & 4 & & \\ \hline 5 & & & \\ \hline 8 & & & \\ \hline \end{array}$$

These tableaux contain eight numbers in four rows, so we try to recover a four-phased construction and annotate the x and y coordinates of an 8×8 grid with the index of the phases at which we wish a corresponding shadowlines to leave the picture.



We claim there is a unique way to draw the shadowlines of phases 4, 3, 2 and 1 (one after the other) that correspond to a forward construction (see Figure 26).

First connect the indices of the last phase with L -shaped shadowlines (the lines of the last phase do not have convex bends). In our example, there is only one such line (drawn in red). Now draw the shadowlines for phase 3, again there is only one in our example. It must connect the corresponding numbers at the border of the picture and have a concave bend at the convex bend of the shadowline of phase 4 on the way, there is only one way to do it. The lines for phase 2 (shown in blue) are added one after the other (right-most first). The first line must visit the convex bend at $(7, 5)$ if it didn't, then this point would not be reachable for the next blue shadowline (since they must not cross). In the last step the shadowlines of phase 1 are added (yellow). Convince yourself that there is, again, only one way to draw them: Every yellow shadowline must visit exactly those convex bends of the blue lines that were not visited by a previous yellow shadowline and that would otherwise become unreachable for subsequent yellow shadowlines.

Now the permutation can be obtained by taking the convex bends of the shadowlines of phase 1. It is: $\pi = 28561437$.

With techniques similar to what we have already seen one can show that this backwards procedure works for arbitrary tableaux. \square

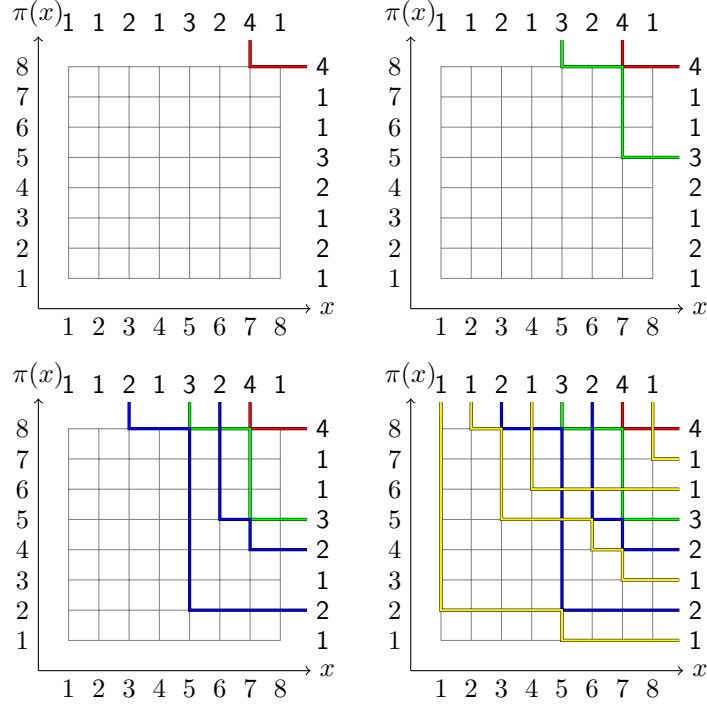
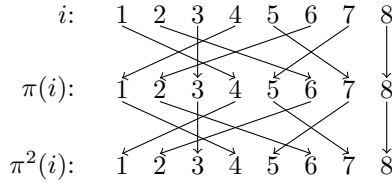


Figure 26: Reverse construction of the Robinson-Schensted-Correspondence.

4.3.1 Counting Tableaux

Given the geometric variant of the Robinson-Schensted-Correspondence (the original construction, although equivalent, was formulated very differently), it is not hard to count all Young tableaux with n elements.

We say a permutation $\pi \in S_n$ is an *involution* if $\pi^2 = \text{id}$, i.e. applying π twice yields the identity permutation. This is equivalent to saying that the disjoint cycle decomposition of π contains only cycles of size 2 (transpositions) and cycles of size 1 (fixed points). For instance:



We would write this involution as $\pi = (1\ 4)(2\ 6)(7\ 5)$.

Lemma 4.11. Let $\pi \xrightarrow{RSC} (T_1, T_2)$, i.e. π is mapped to (T_1, T_2) via the Robinson-Schensted-Correspondence. Then

(i) $\pi^{-1} \xrightarrow{RSC} (T_2, T_1)$.

(ii) π is an involution iff $T_1 = T_2$.

(iii) The length of the longest increasing subsequence in π is the length of the first row in T_1 .

Proof. (i) Recall that the construction starts with

$$X(\pi) = \{(i, \pi(i)) \mid i \in [n]\} = \{(\pi^{-1}(j), j) \mid j \in [n]\}.$$

The point set $X(\pi^{-1})$ is therefore obtained by flipping the point set $X(\pi)$ along the diagonal $x = y$, in other words, changing the roles of x and y . The shadowlines we obtain will also be flipped, so every x -coordinate we would have written into T_1 is now a y -coordinate and therefore written into T_2 and vice versa. This just means that the roles of T_1 and T_2 are swapped.

(ii) $\pi^2 = \text{id} \Leftrightarrow \pi = \pi^{-1} \stackrel{(i)}{\Leftrightarrow} (T_1, T_2) = (T_2, T_1) \Leftrightarrow T_1 = T_2$.

(iii) Consider an increasing subsequence, i.e.

$$i_1 < i_2 < \dots < i_k \text{ with } \pi(i_1) < \pi(i_2) < \dots < \pi(i_k).$$

The corresponding set $Y = \{(i_j, \pi(i_j)) \mid j \in [k]\}$ is a chain in $X(\pi)$. Now we can use a sub claim from the last theorem which asserted that the length of the i -th row of the tableaux is the length of a longest chain in X_i , using it for $i = 1$.

□

Note that by Lemma 4.11(ii) we have that the number i_n of involutions of $[n]$ and the number of standard Young tableaux with n squares coincide. The following Lemma therefore gives a recurrence for the number of Young tableaux.

Lemma 4.12. *For the number i_n of involutions of $[n]$ we have*

$$i_1 = 1, \quad i_2 = 2, \quad i_n = i_{n-1} + (n-1)i_{n-2}.$$

Proof. Consider involutions in their disjoint cycle decomposition. The cycles contain one or two elements, so think of an involution as an arrangement of n labeled balls in unlabeled boxes and one or two balls per box. To count them, we distinguish two cases:

Case 1: The ball with label n is in its own box. The rest is an arrangement with $n-1$ balls.

Case 2: The ball with label n is in a box together with another ball x . There are $n-1$ choices for x and the rest is an arrangement with $n-2$ balls.

This proves the claim.

□

4.3.2 Counting Tableaux of the Same Shape

Let $\mathcal{T}(\lambda)$ be the set of all standard Young tableaux with shape λ . Take for instance

$$\lambda = \begin{array}{|c|c|c|} \hline & & \\ \hline & & \\ \hline \end{array}, \quad \mathcal{T}(\lambda) = \left\{ \begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 4 & & 5 \\ \hline \end{array}, \begin{array}{|c|c|c|} \hline 1 & 2 & 4 \\ \hline 3 & & 5 \\ \hline \end{array}, \begin{array}{|c|c|c|} \hline 1 & 2 & 5 \\ \hline 3 & & 4 \\ \hline \end{array}, \begin{array}{|c|c|c|} \hline 1 & 3 & 4 \\ \hline 2 & & 5 \\ \hline \end{array}, \begin{array}{|c|c|c|} \hline 1 & 3 & 5 \\ \hline 2 & & 4 \\ \hline \end{array} \right\}.$$

In general, let $\lambda = (n_1, n_2, \dots, n_m)$ where $n_1 \geq n_2 \geq \dots \geq n_m$ are numbers that sum up to n . With this in mind we define the function $t : \mathbb{Z}^m \rightarrow \mathbb{N}$ as

$$t(n_1, \dots, n_m) = \begin{cases} |\mathcal{T}((n_1, \dots, n_m))| & \text{if } n_1 \geq n_2 \geq \dots \geq n_m \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

For technical reasons, we allow trailing zero-length rows, for instance we could write $\mathcal{T}((3, 2, 0, 0)) = \mathcal{T}((3, 2, 0)) = \mathcal{T}((3, 2)) = 5$. We claim the function is fully characterized by the following identities:

- (1) If the numbers n_i fail to be weakly decreasing then

$$t(n_1, \dots, n_m) = 0.$$

- (2) If the last number is zero, then it can be removed

$$t(n_1, \dots, n_m, 0) = t(n_1, \dots, n_m).$$

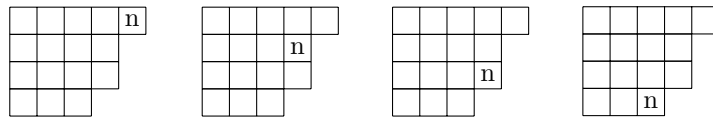
- (3) If the numbers are weakly decreasing and none is zero then

$$\begin{aligned} t(n_1, \dots, n_m) &= \sum_{i=1}^m t(n_1, n_2, \dots, n_{i-1}, n_i - 1, n_{i+1}, \dots, n_m) \\ &= t(n_1 - 1, \dots, n_m) + t(n_1, n_2 - 1, \dots, n_m) + \dots \end{aligned}$$

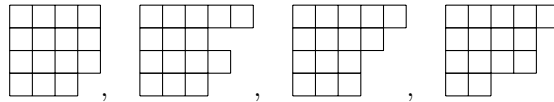
- (4) If only one number $n \geq 0$ is left then

$$t(n) = 1.$$

It is obvious that t fulfills (1), (2) and (4). And (3) is just the observation that the biggest number n must be the last number in one of the m rows. Consider for instance $\lambda = (5, 4, 4, 3)$ then in any Young tableau the number $n = 16$ will be in one of the following positions:



For each case, we count the number of Young diagrams of the shape where the corresponding square was removed, i.e. we consider the shapes



So we compute $t(4, 4, 4, 3) + t(5, 3, 4, 3) + t(5, 4, 3, 3) + t(5, 4, 4, 2)$. You may object that the second shape is not valid, but that's no problem since $t(5, 3, 4, 3)$ was defined to be zero.

Note that (1)-(4) gives a complete recursion, i.e. for each input the value is either defined explicitly or given in terms of values for smaller inputs (in (2) the number of terms is decreased, in (3) the sum of the inputs is decreased). This

means there is a unique solution. While we do not try the long and arduous journey of discovering the solution ourselves, we can, given the solution, verify that it is indeed correct.

To this end, we need to first examine the *Vandermonde determinant* which is defined⁷ as

$$\Delta(x_1, \dots, x_m) := \prod_{1 \leq i < j \leq m} (x_i - x_j).$$

It has the curious property that swapping two input values changes the sign of the output. To see this, observe firstly what happens if the adjacent inputs x_i and x_{i+1} in the argument list of Δ are swapped for some $i \in [m-1]$. All factors remain unchanged with the exception of $(x_i - x_{i+1})$ which is replaced by $(x_{i+1} - x_i) = -(x_i - x_{i+1})$ as claimed. If two non-adjacent values x_i, x_{i+k} are swapped, then this can be simulated by an odd number of swaps of adjacent elements. Think of a race with n runners where Alice is in place i and Bob in place $i+k$. We can make them change places as follows: First Bob overtakes k people, putting him in front of Alice and then Alice (who is now in position $i+1$) must fall behind $k-1$ positions. This gives $2k-1$ overtaking operations in total, an odd number.

We remark (without proof), that the Vandermonde is indeed a determinant, namely

$$\Delta(x_1, \dots, x_m) = (-1)^{\binom{m}{2}} \det \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{m-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{m-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^{m-1} \end{pmatrix}.$$

Before we can get back to counting tableaux, we need to prove a technical Lemma:

Lemma 4.13. *We have the following identity on polynomials over x_1, x_2, \dots, x_m, y :*

$$\sum_{k=1}^m x_k \Delta(x_1, \dots, x_k + y, \dots, x_m) = (x_1 + x_2 + \dots + x_m + \binom{m}{2} y) \Delta(x_1, \dots, x_m).$$

Proof. Let g be the polynomial given as the sum on the left hand side. Observe what happens if we swap the roles of x_i and x_j (with $i < j$) in g . The summands for $k \notin \{i, j\}$ will just change sign (by our previous observation). The summand with $k = i$ turns into:

$$x_j \Delta(x_1, \dots, \underset{\substack{\uparrow \\ \text{position } i}}{x_j + y}, \dots, \underset{\substack{\uparrow \\ \text{position } j}}{x_i}, \dots, x_m) = -x_j \Delta(x_1, \dots, \underset{\substack{\uparrow \\ \text{position } i}}{x_i}, \dots, \underset{\substack{\uparrow \\ \text{position } j}}{x_j + y}, \dots, x_m).$$

So the term for $k = i$ became the negated term with $k = j$, and this works vice versa as well — altogether, the value of g changes sign if x_i and x_j swap roles. Now consider the case of $x_i = x_j$, then swapping x_i and x_j obviously *doesn't* change anything. The only number that doesn't change if its sign changes is zero, so $g = 0$ whenever two x_i and x_j coincide (for $i \neq j$).

Now think of all the variables x_2, \dots, x_m, y as some (distinct) integer constants and of x_1 as the only actual variable, then g is a polynomial of degree

⁷Our definition differs from the usual one in that it may have a different sign.

n over x_1 . It has several zeroes, one of which is $x_1 = x_2$. This allows us to divide g by the degree 1 polynomial $x_1 - x_2$ (using polynomial long division) and we obtain $g = p \cdot (x_1 - x_2)$ for some polynomial p . In the same way we separate the other zeroes getting $g = p' \cdot (x_1 - x_2)(x_1 - x_3) \dots (x_1 - x_m)$ for some polynomial p' . Then, looking at p' , we switch perspective thinking of x_2 as the variable and of the other values as (suitable distinct) constants. Looking at it that way we find the zeroes $x_2 = x_3, x_2 = x_4, \dots, x_2 = x_m$ of p' and can separate corresponding factors from p' . We repeat this for the remaining $m - 2$ variables as well. With this we eventually obtain:

$$g = \hat{p} \cdot \prod_{1 \leq i < j \leq m} (x_i - x_j). \quad (\star)$$

for some polynomial \hat{p} . Note that this was the important step. Since this is not an algebra lecture, we allowed ourselves to be a bit sketchy: Formally we would have to argue that no funny business is going on when doing the polynomial long division and switching perspective between thinking of the x_i as constants or variables.

Given (\star) , everything falls into place: The degree of the polynomial g is $\binom{m}{2} + 1$ while the degree of the right hand side is $\binom{m}{2}$ plus the degree of \hat{p} . So the degree of \hat{p} is one!

This means that g is of the form:

$$\begin{aligned} g &= (a_1x_1 + a_2x_2 + \dots + a_mx_m + by) \cdot \prod_{1 \leq i < j \leq m} (x_i - x_j) \\ &= (a_1x_1 + a_2x_2 + \dots + a_mx_m + by)\Delta(x_1, \dots, x_m). \end{aligned}$$

for some constants a_1, \dots, a_m, b still to be determined⁸. But recall that g was originally defined to be:

$$g = \sum_{k=1}^m x_k \Delta(x_1, \dots, x_k + y, \dots, x_m).$$

Since the equation must hold for the special case of $y = 0$, we quickly see that $a_1 = a_2 = \dots = a_m = 1$.

Now to determine b , we multiply out g and collect the sum of all monomials containing y with multiplicity 1. We start by analyzing the k -th summand of g :

$$\begin{aligned} &x_k \cdot \Delta(x_1, \dots, x_k + y, \dots, x_m) \\ &= x_k \cdot \left(\prod_{\substack{1 \leq i < j \leq m \\ i, j \neq k}} (x_i - x_j) \right) (x_1 - x_k - y)(x_2 - x_k - y) \dots (x_k + y - x_m). \end{aligned}$$

To get monomials with a single y , choose y from one of the factors and choose the part without y everywhere else. This allows to reassemble Δ with the exception of a single missing factor (we write it into the denominator). The

⁸You may object that \hat{p} may have a constant term. We call a polynomial *homogeneous*, if all monomials have full degree. Since g and $\Delta(x_1, \dots, x_m)$ are homogeneous, we can conclude that \hat{p} must be homogeneous as well.

$m - 1$ summands (one for each occurrence of y) add up to

$$\begin{aligned}
& x_k \cdot \left(\prod_{1 \leq i < j \leq m} (x_i - x_j) \right) \left(\frac{-y}{x_1 - x_k} + \frac{-y}{x_2 - x_k} + \dots + \frac{y}{x_k - x_m} \right) \\
&= y \Delta(x_1, \dots, x_m) \left(- \sum_{i=1}^{k-1} \frac{x_k}{x_i - x_k} + \sum_{i=k+1}^m \frac{x_k}{x_k - x_i} \right) \\
&= y \Delta(x_1, \dots, x_m) \sum_{i \neq k} \frac{-x_k}{x_i - x_k}.
\end{aligned}$$

If we sum up these terms for all k we get

$$\begin{aligned}
& y \Delta(x_1, \dots, x_m) \sum_{k=1}^m \sum_{i \neq k} \frac{-x_k}{x_i - x_k} \\
&= y \Delta(x_1, \dots, x_m) \sum_{1 \leq i < k \leq m} \left(\frac{-x_k}{x_i - x_k} + \frac{-x_i}{x_k - x_i} \right) \\
&= y \Delta(x_1, \dots, x_m) \sum_{1 \leq i < k \leq m} \frac{x_k - x_i}{x_k - x_i} = y \Delta(x_1, \dots, x_m) \binom{m}{2}.
\end{aligned}$$

This shows $b = \binom{m}{2}$, completing the proof. \square

Theorem 4.14.

$$t(n_1, \dots, n_m) = \frac{\Delta(x_1, \dots, x_m) n!}{x_1! x_2! \dots x_m!}$$

where $n = \sum_{i=1}^m n_i$, $x_i = n_i + m - i$ for $i = 1, \dots, m$ and $x_1 \geq \dots \geq x_m \geq 0$.

Proof. We show that the right hand side is a solution to the recurrence we found for t , i.e. if we define t by the right hand side then (1) – (4) from page 101 hold. First note what happens if x_i are not strictly decreasing, i.e. $x_i = x_{i+1}$ for some i . This gives a value of 0 since the factor $0 = x_i - x_{i+1}$ occurs in $\Delta(x_1, \dots, x_m)$. We also know:

$$n_i + m - i = n_{i+1} + m - (i + 1) \Leftrightarrow n_i = n_{i+1} - 1,$$

so the n_i are not weakly decreasing which means these values do not correspond to a valid shape for Young tableaux. Some other invalid shapes were already excluded by restricting ourselves to weakly decreasing x_i . All in all, we are consistent with:

$$(1) \quad t(n_1, \dots, n_m) = 0 \text{ unless } n_1 \geq \dots \geq n_m \geq 0.$$

The second thing we need to show is

$$(2) \quad t(n_1, \dots, n_{m-1}, 0) = t(n_1, \dots, n_{m-1}).$$

Since $n_m = 0$ means $x_m = 0$, we calculate:

$$\begin{aligned}
\Delta(x_1, \dots, x_{m-1}, 0) &= \prod_{1 \leq i < j \leq m} (x_i - x_j) = \prod_{1 \leq i < j \leq m-1} (x_i - x_j) \prod_{i=1}^{m-1} x_i \\
&= \Delta(x_1, \dots, x_{m-1}) \prod_{i=1}^{m-1} x_i = \Delta(x_1 - 1, \dots, x_{m-1} - 1) \prod_{i=1}^{m-1} x_i.
\end{aligned}$$

With this it is easy to verify:

$$\begin{aligned} t(n_1, \dots, n_{m-1}, 0) &= \frac{\Delta(x_1, \dots, x_{m-1}, 0)n!}{x_1! \dots x_{m-1}!0!} \\ &= \frac{\Delta(x_1 - 1, \dots, x_{m-1} - 1)n!}{(x_1 - 1)! \dots (x_{m-1} - 1)!} = t(n_1, \dots, n_{m-1}). \end{aligned}$$

The hard part is (3), but we did most of the work already in the last Lemma.

$$(3) \quad t(n_1, \dots, n_m) = \sum_{k=1}^m t(n_1, \dots, n_k - 1, \dots, n_m).$$

First note:

$$\sum_{i=1}^m x_i = \sum_{i=1}^m (n_i + m - i) = n + \sum_{j=0}^{m-1} j = n + \binom{m}{2}.$$

Now Lemma 4.13 gives for $y = -1$

$$\sum_{k=1}^m x_k \Delta(x_1, \dots, x_k - 1, \dots, x_m) = \underbrace{(x_1 + \dots + x_m - \binom{m}{2})}_n \Delta(x_1, \dots, x_m)$$

Hence:

$$\begin{aligned} t(n_1, \dots, n_m) &= \frac{\Delta(x_1, \dots, x_m)n!}{x_1! \dots x_m!} = \frac{\sum_{k=1}^m x_k \Delta(x_1, \dots, x_k - 1, \dots, x_m)(n-1)!}{x_1! \dots x_m!} \\ &= \sum_{k=1}^m \frac{\Delta(x_1, \dots, x_k - 1, \dots, x_m)(n-1)!}{x_1! \dots (x_k - 1)! \dots x_m!} = \sum_{k=1}^m t(n_1, \dots, n_k - 1, \dots, n_m) \end{aligned}$$

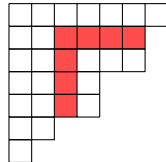
The last property is again easy to verify. For $m = 1$ and $n = n_1$, we have:

$$(4) \quad t(n_1) = \frac{\Delta(x_1)n!}{x_1!} = \frac{\Delta(n_1)n_1!}{n_1!} = \Delta(n_1) = 1.$$

where $\Delta(n_1) = 1$ because it is an empty product. \square

With the last Theorem it would already be possible to count Young tableaux of any shape fairly conveniently. However, with a geometric interpretation, it gets even easier. This geometric interpretation uses *hooks*.

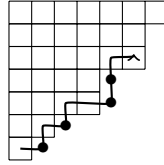
Consider a partition $\lambda = (n_1, \dots, n_m)$ and a square $(i, j) \in \lambda$, by which we mean the square in row i and column j . The *hook* $h_{i,j}$ rooted in (i, j) is the union of all squares to the right of the square in the same row and below of the square in the same column. Here is a picture, a Ferrer diagram where the hook $h_{2,3}$ is highlighted, it has length 7.



Theorem 4.15 (Hook length formula). *Let λ be a partition of n . Then*

$$t(\lambda) = \frac{n!}{\prod_{(i,j) \in \lambda} |h_{i,j}|}.$$

Proof. Let $\lambda = (n_1, \dots, n_m)$. We multiply the lengths of all the hooks, going through them row by row. For row i , the first hook $h_{i,1}$ starts at square $(m, 1)$ then goes upwards and rightwards and ends at square (i, n_i) . It has length $(n_i + m - i) = x_i$. The other hooks in row i also end in (i, n_i) but they start in other positions. We go through these positions from left to right as shown in the following figure (for $i = 3$).

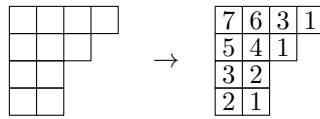


Now if there was a hook starting in each integer position this line passes through, the product of the hook lengths would just be $x_i!$. However, the positions marked with a dot do not correspond to starts of hooks so for each dot we have to divide by the length that a hook starting there would have. The dot in line j ($j > i$) is in position $(j, n_j + 1)$ so a hook going to (i, n_i) has length $(j - i + n_i - n_j) = (x_i - x_j)$. So the product of all valid hooks for line i is $\frac{x_i!}{\prod_{j>i} (x_i - x_j)}$. Now for the product of all hook lengths of all rows we get:

$$\prod_{(i,j) \in \lambda} |h_{i,j}| = \prod_{1 \leq i \leq m} \frac{x_i!}{\prod_{j>i} (x_i - x_j)} = \frac{x_1! \dots x_m!}{\prod_{i<j} (x_i - x_j)} = \frac{x_1! \dots x_m!}{\Delta(x_1, \dots, x_m)} \stackrel{4.14}{=} \frac{n!}{t(\lambda)}.$$

□

Take for instance the following Ferrer diagram, that we annotate with the lengths of the hooks rooted at the respective position:



The number of Young tableaux of this shape is by the Hook length formula $11!$ divided by all the hook lengths, so:

$$\frac{11!}{7 \cdot 6 \cdot 3 \cdot 1 \cdot 5 \cdot 4 \cdot 1 \cdot 3 \cdot 2 \cdot 2 \cdot 1} = \frac{11 \cdot 10 \cdot 9 \cdot 8}{3 \cdot 2} = 11 \cdot 10 \cdot 3 \cdot 4 = 1320.$$

5 Partially Ordered Sets

Example 5.1. As was accurately observed by Randall Munroe, creator of [xkcd](https://xkcd.com/388/), different fruit not only differ in their tastiness, but also in the difficulty of preparing them. Some types of fruit are clearly superior to others, for instance

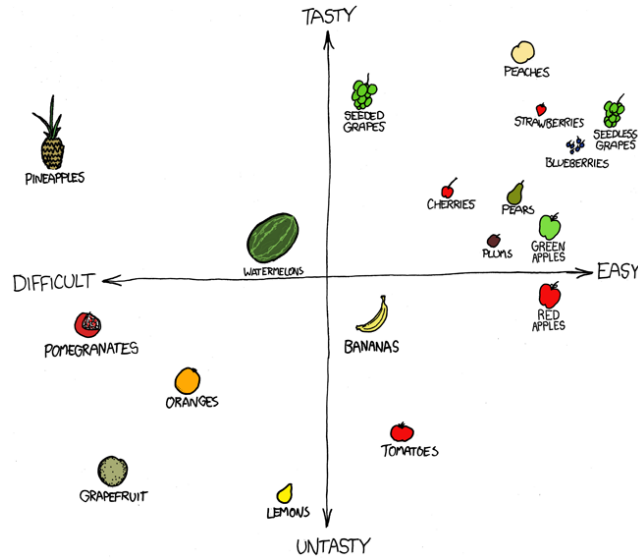


Figure 27: <https://xkcd.com/388/>: Coconuts are so far down to the left they couldn't be fit on the chart. Ever spent half an hour trying to open a coconut with a rock? – Randall Munroe

seedless grapes are both more tasty and easier to prepare than oranges. However, if you compare pineapples with bananas then pineapples are more tasty but harder to prepare, so there is no clear winner.

To model these situations where some things are bigger/better/above/dominating others things but some pairs of things may also be incomparable/on the same level/of equal rank, we introduce the concept of partial orders.

Definition 5.2. A *partially ordered set* or *poset* or *partial order* is a pair $P = (X, \leq)$ where X is a (for us usually finite) set and \leq is a binary relation that is:

- (i) reflexive: $x \leq x$, $\forall x \in X$.
- (ii) transitive: $(x \leq y \text{ and } y \leq z) \Rightarrow x \leq z$, $\forall x, y, z \in X$.
- (iii) antisymmetric: $(x \leq y \text{ and } y \leq x) \Rightarrow x = y$, $\forall x, y \in X$.

You probably already know some (partial) orders, and recognize a few examples from Table 3. Note that a *total order* (where for any $x, y \in X$ we have $x \leq y$ or $y \leq x$) is a special case of a partial order. We now introduce some notation.

- If $x \leq y$ and $x \neq y$ then we write $x < y$.
- If $x \leq y$ we also write $y \geq x$.

X	\leq	Example Relations
natural numbers	order by value	$23 \leq 42, 42 \not\leq 23$
natural numbers	order by divisibility	$6 \not\leq 7, 7 \not\leq 6, 13 \leq 91$
polynomials	order by divisibility	$X - 7 \leq X^2 - 10X + 21$
words over $\{A, \dots, Z\}$	order lexicographically	ELEPHANT \leq MOUSE
real intervals	order by inclusion	$[3, 5] \leq [2, 7], [2, 3] \not\leq [4, 6]$
real intervals	“completely left of”	$[3, 5] \not\leq [2, 7], [2, 3] \leq [4, 6]$
real valued functions	point-wise domination	$1 + x \not\leq x^2, \sin(x) \leq 2$
subsets of $[N]$	inclusion	$\{1, 3\} \leq \{1, 2, 3, 5\}$

Table 3: Some posets $P = (X, \leq)$ with informal definitions of their order relations and examples.

- If neither $x \leq y$ nor $y \leq x$ then x, y are *incomparable* denoted by $x \parallel y$.
- If $x < y$ and there is no z with $x < z < y$, then $x < y$ is a *cover relation* (or *cover* for short), denoted by $x \lessdot y$.

We depict a poset by its *Hasse diagram*. In it, every $x \in X$ corresponds to a point in the plane and for every $x, y \in X$ with $x < y$ the corresponding points are connected by a y -monotone line where x is the lower (with smaller y -coordinate) endpoint. In particular “transitive edges” are omitted meaning $x, y \in P$ are *not* directly connected if there is a z such that $x < z < y$. The poset from Example 5.1 has the Hasse diagram shown in Figure 28.

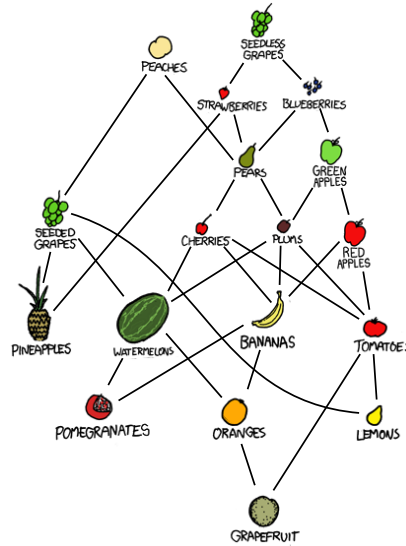


Figure 28: Hasse Diagram for the poset originating from Figure 5.1. For instance lemons and bananas are connected by a line since bananas are both tastier and easier to prepare. Lemons and cherries are not connected by a direct line since they are already connected via bananas. Bananas and watermelons are not connected at all since they are incomparable.

Sometimes it may be easier to define the cover relations than to define the entire relation.

Example. Let X be the set of states of a Rubik's Cube. For $x \in X$ define $r(x)$ to be the minimum number of moves needed to solve the cube from state x .

We want $x \leq y$ if and only if $r(x) < r(y)$ and x and y can be transformed into one another by one move. This is the cover relation of a poset (without proof).

We now define some key concepts and corresponding notation.

Definition 5.3. Let $P = (X, \leq)$ poset.

- A *chain* is a sequence of elements in increasing order, i.e. $y_1 < y_2 < \dots < y_k$ for $y_i \in X$. It has *length* k .
- The *height* of P is the length of a longest chain in P .
- A set $Y \subseteq X$ where elements are pairwise incomparable ($y_1 \parallel y_2$ for all $y_1 \neq y_2 \in Y$) is an *antichain*.
- The *width* of P is the size of a largest antichain in P .
- The sets of *minimal* and *maximal* elements of P are given as

$$\begin{aligned}\min(P) &:= \{x \in X \mid \forall y \in X : x \leq y \text{ or } x \parallel y\}, \\ \max(P) &:= \{x \in X \mid \forall y \in X : x \geq y \text{ or } x \parallel y\}.\end{aligned}$$

Looking at the poset P from Figure 28 again, an example for a chain would be {Watermelons, Pears, Strawberries}. There are several longest chains, one of which is {Grapefruit, Oranges, Bananas, Plums, Pears, Blueberries, Seedless Grapes}. The height of P is therefore 7. An example for an antichain is {Pineapple, Cherries, Plums, Red Apples}. No other antichain is longer so the width of P is 4. The maximal elements are $\max(P) = \{\text{Seedless Grapes, Peaches}\}$ and the minimal elements are $\min(P) = \{\text{Pineapples, Pomegranates, Grapefruit, Lemons}\}$. We now study partitions of posets into chains and antichains. We start with the easier case.

Theorem 5.4 (Antichain Partitioning). *The elements of every poset $P = (X, \leq)$ can be partitioned into $h(P)$ antichains (and not less).*

Proof. Note first that we cannot partition P into fewer antichains: No antichain can contain two elements of a chain, since elements of chains are pairwise comparable and elements of antichains are pairwise incomparable. Since P contains a chain Y of size $h(P)$, at least $h(P)$ antichains are needed.

To see that $h(P)$ antichains suffice, first note that $\min(X)$ is an antichain: Two minimal elements are always unrelated, otherwise the “bigger” minimal element would not be minimal at all. Also, every maximal chain in P contains an element from $\min(X)$: If $Y = \{y_1, \dots, y_k\}$ is a maximal chain with $y_1 < \dots < y_k$ and y_1 is not minimal, then we would find $y_0 < y_1$ and therefore a longer chain.

So we take the first antichain to be $\min(P)$, then we still need to partition $X \setminus \min(P)$ into $h(P) - 1$ antichains. Since the maximal chains in $X \setminus \min(P)$ have size at most $h(P) - 1$ we can do this by induction. \square

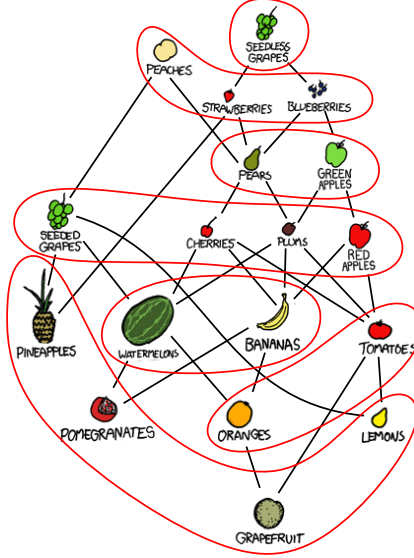


Figure 29: Partition of the poset into 7 antichains obtained by iteratively putting the minimal remaining elements into a new antichain.

To get back to our fruit example, Figure 29 shows how the poset can be partitioned into 7 antichains.

Theorem 5.5 (Dilworth's Theorem). *Every poset $P = (X, \leq)$ can be partitioned into $w(P)$ chains (and not less).*

Proof. Clearly, $w(P)$ chains are necessary: P contains an antichain of size $w(P)$ and no two of its elements can be contained in the same chain.

To show $w(P)$ chains suffice, we do induction on $|X|$. Consider a maximum antichain $A = \{x_1, \dots, x_{w(P)}\}$.

The idea is to split P along A into two parts, take for instance our fruit poset and $A = \{\text{Seeded Grapes, Cherries, Plums, Red Apples}\}$, then the two parts are shown in Figure 30.

Formally we define $P_1 = (X_1, \leq)$ and $P_2 = (X_2, \leq)$ with elements

$$X_1 := \{y \in X \mid \exists x \in A: x \leq y\}, X_2 := \{y \in X \mid \exists x \in A: y \leq x\},$$

and the same relations as before. Note that with this definition $X_1 \cap X_2 = A$, since if $y \in (X_1 \cap X_2)$ then there is $x_1 \in A$ and $x_2 \in A$ with $x_1 \leq y \leq x_2$. Since A is an antichain this implies $x_1 = x_2 = y$ so $y \in A$.

Now if we can partition P_1 into chains $C_1^1, \dots, C_{|A|}^1$ and P_2 into chains $C_1^2, \dots, C_{|A|}^2$ where C_i^1 and C_i^2 both contains x_i then we can attach the chains to one another obtaining a partition of P into chains $C_1, \dots, C_{|A|}$. We can find these partitions by induction *unless* P_1 or P_2 fail to be smaller than P .

Convince yourself that $P_1 = P \Leftrightarrow A = \min(P)$ and $P_2 = P \Leftrightarrow A = \max(P)$. So our only problem is the case where there is no largest antichain except for $\min(P)$, $\max(P)$ or both.

In that case, let C be any maximal chain. In the same way as in the previous Theorem we argue that $\min(P) \cap C \neq \emptyset$ and $\max(P) \cap C \neq \emptyset$. Then $P \setminus C$ has

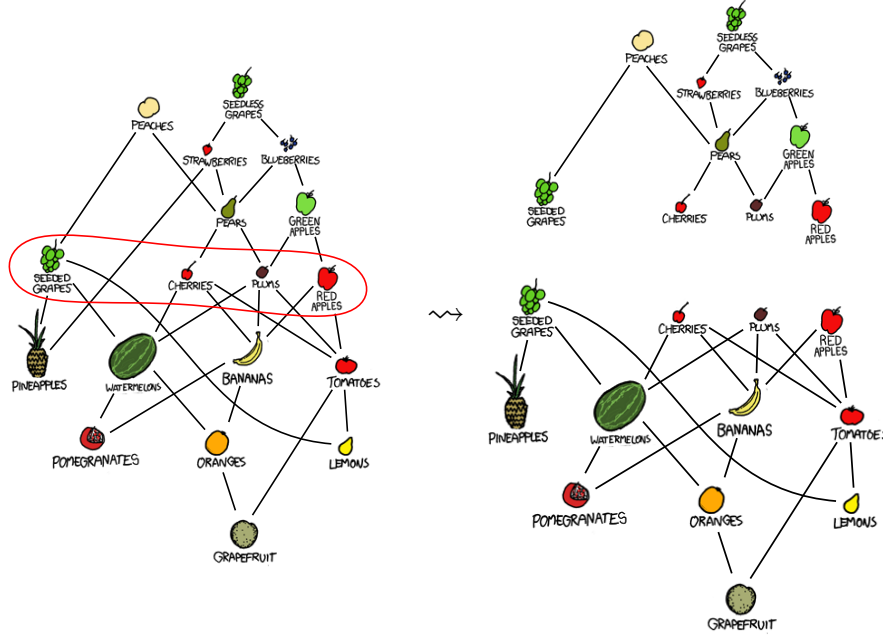


Figure 30: We split the poset P along an antichain into P_1 (the upper half) and P_2 (the lower half). The elements of the antichain are contained in both posets after the split.

width $w(P) - 1$ (since we removed an element from each largest antichain) so it can be partitioned into $w(P) - 1$ chains by induction and we are done. \square

5.1 Subposets, Extensions and Dimension

We now define when a poset is part of another poset, that is, when $P = (X, \leq_P)$ is a subposet of $Q = (Y, \leq_Q)$. This should be the case if $X \subseteq Y$ and $\forall x_1, x_2 \in X : (x_1 \leq_P x_2 \Leftrightarrow x_1 \leq_Q x_2)$. In that case we say P is the *subposet of Q induced by X* . But since we do not care about relabeling of elements, the notion of subposet is actually a bit more generous:

P is a *subposet* of Q if there is an injective function $f : X \rightarrow Y$ such that $\forall x_1, x_2 \in X : (x_1 \leq_P x_2 \Leftrightarrow f(x_1) \leq_Q f(x_2))$. The function f is an *embedding* and the image of f is a *copy* of P in Q .

We say $Q = (X, \leq_Q)$ is an *extension* of $P = (X, \leq_P)$ (note that they use the same ground set X) if $\forall x_1, x_2 \in X : x_1 \leq_P x_2 \Rightarrow x_1 \leq_Q x_2$. When interpreting relations as subsets of $X \times X$ we could write $\leq_P \subseteq \leq_Q$.

We say $L = (X, \leq_L)$ is a *linear extension* of P if it is an extension of P and a total order (meaning no two elements are incomparable in \leq_L).

Lemma 5.6. *Let $P_1 = (X, \leq_1), P_2 = (X, \leq_2)$ be two posets on the same ground set X . Then we define $P = P_1 \cap P_2 := (X, \leq)$ with the relation $\leq := \leq_1 \cap \leq_2$, meaning*

$$x \leq y \Leftrightarrow (x \leq_1 y \text{ and } x \leq_2 y).$$

We claim P is a poset and P_1 and P_2 are extensions of P .

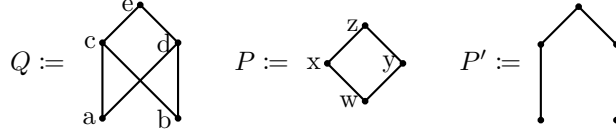


Figure 31: With the posets as shown (via their Hasse diagrams) P is a subposet of Q as witnessed by the map $w \mapsto a, x \mapsto c, y \mapsto d, z \mapsto e$. Actually, there are four different embeddings, since w could also be mapped to b and x and y can also be mapped to c and d the other way round. These two maps correspond to two copies of P in Q , namely $\{a, c, d, e\}$ and $\{b, c, d, e\}$. The poset P' is *not* a subposet of Q . However, Q is an extension of P' .

Proof. It is easy to verify that \leq is an order relation. It is also clear that all pairs that are related via \leq are related via \leq_1 and \leq_2 , so \leq_1 and \leq_2 are extensions of \leq . \square

Two-dimensional posets. Recall the poset P from the fruit example. It has two natural linear extensions: The first is the total order L_{Taste} in which the fruit are arranged in order of increasing tastiness:

$$\text{Lemons} \leq_{L_{\text{Taste}}} \text{Grapefruit} \leq_{L_{\text{Taste}}} \text{Tomatoes} \leq_{L_{\text{Taste}}} \dots \leq_{L_{\text{Taste}}} \text{Peaches}.$$

In other words, L_{Taste} is the order obtained when projecting all elements to the tastiness-axis (note how this is an extension of P and a total order, assuming no two fruit have exactly the same tastiness) The second is the total order L_{Ease} , in which the fruits are arranged in order of increasing ease of preparing them:

$$\text{Pineapples} \leq_{L_{\text{Ease}}} \text{Pomegranates} \leq_{L_{\text{Ease}}} \dots \leq_{L_{\text{Ease}}} \text{Seedless Grapes}.$$

Fruit are ordered if and only if they are ordered according to both linear orders so $P = L_1 \cap L_2$.

We want to capture this property in a new notion and define: A poset is *2-dimensional* if it has a two dimensional picture⁹, i.e. an assignment $f : P \rightarrow \mathbb{R}^2$ of positions in the plane to each element such that $x < y$ in P if and only if $f(y)$ is above and right of $f(x)$.

Not every poset is 2-dimensional. Consider the spider poset with three legs, the Hasse Diagram is shown on the left of Figure 32. This spider has a head H three knees K_1, K_2, K_3 considered bigger than the head and three feet F_1, F_2, F_3 considered smaller than the corresponding knee. There are no further relations.

We try to find a two dimensional embedding into the plane, i.e. assign positions in the plane to each element (see right side of Figure 32).

The head H has to go somewhere which partitions the plane into four quadrants (above and right of H , below and right of H , above and left of H , below and left of H). The knees must all go above and right of H since they are bigger than H . Since knees are incomparable, the ones with bigger x coordinate must

⁹Strictly speaking we would have to say: It has a two-dimensional picture but no one-dimensional picture, for details refer to the general definition later.

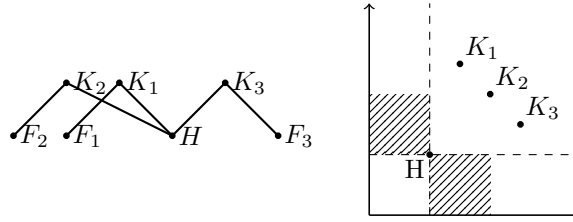
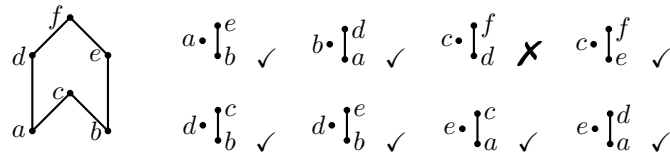


Figure 32: On the left: The spider with three legs. On the right: Sketch for the proof that the spider is not two-dimensional.

have smaller y coordinate so the knees must lie on a decreasing curve as shown. Without loss of generality, K_2 is the second point on this decreasing curve. Now observe that there is no suitable point to put F_2 : It cannot be in the quadrant below and left of H nor in the quadrant above and right of H since $H \parallel F_2$. Since it must also be below and left of K_2 , it must be in the shaded area. But points in that area are below and left of either K_1 or K_3 (or both), which is not allowed since $F_2 \parallel K_1$ and $F_2 \parallel K_3$. This completes the proof.

Sums of Chains. We define an additional bit of notation. By $i \oplus k$ for $i, k \in \mathbb{N}$ we mean the poset that is the union of a chain of length i and a chain of length k and no comparabilities between the chains. For instance $1 \oplus 1$ is an antichain of size 2. The poset important in the following is $1 \oplus 2 = \cdot \downarrow$, let's label with $a \cdot \downarrow_b^c$. It consists of three elements a, b, c with $b < c$ and $a \parallel b, a \parallel c$. Note that $\cdot \downarrow$ has three linear extensions, $a < b < c$, $b < c < a$, $b < a < c$, the last of which will have a special standing in the following: Given a linear extension L of some poset P and a copy of $\cdot \downarrow$ in P then we say L *orders* this copy of $\cdot \downarrow$ if the lonely element (corresponding to a above) does *not* come in between the other two elements (corresponding to b and c above). Take for instance the linear ordering $L : a < d < b < c < e < f$ of following poset (shown on the left). It contains 8 copies of $\cdot \downarrow$ listed on the right and almost all of them are ordered by L .

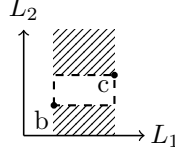


The only exception is $c \cdot \downarrow_d^f$. It is not obvious whether or not a different linear extension would have ordered every copy of $\cdot \downarrow$ (you will have to find that out on the exercise sheet). What we show now is the significance of the existence of such a linear extension.

Theorem 5.7. *A poset P is 2-dimensional if and only if there is a linear extension L of P ordering all copies of $1 \oplus 2$ in P .*

Proof. “ \Rightarrow ” Assume P is 2-dimensional. Then $P = L_1 \cap L_2$ for the two linear extension L_1, L_2 of P corresponding to an embedding into \mathbb{R}^2 . We claim that any of the linear extensions, say L_1 , orders every copy of $\cdot \downarrow$.

Assume not, then we find elements a, b, c in P that form an unordered copy of $\cdot \downarrow$, i.e. we have the situation $a \cdot \downarrow_b^c$ with $b <_{L_1} a <_{L_1} c$. Since $b \leq_P c$ we also have $b \leq_{L_2} c$ so we are in this situation:

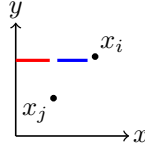


We know from L_1 that, horizontally, a is between b and c , but vertically it must be below b (since $a \parallel b$) and it must be above c (since $a \parallel c$). Clearly this is not possible (it would have to be in both shaded areas at once) which contradicts the assumption that L_1 does not order a copy of $\cdot \downarrow$.

“ \Leftarrow ” Let L be a linear extension of P ordering all copies of $\cdot \downarrow$. From this we construct an embedding into \mathbb{R}^2 . First label the elements of P in accordance with L , i.e. $x_1 <_L \dots <_L x_n$. The x -coordinate of x_i will just be i . We define the y -coordinates from left to right, i.e. we define x_i assuming those of x_1, \dots, x_{i-1} are already given.

Considering the y -coordinates of $\{x_j \mid j < i, x_j \leq_P x_i\}$ and place x_i *barely* above the maximum such y -coordinate.

Clearly, this procedure ensures that if we have $x_j \leq_P x_i$ then x_i is put above and right of x_j . But we also have to ensure that incomparable elements are not placed this way. Assume there is a smallest index i where $x_j \parallel x_i$ for some $j < i$ and the above procedure will still put x_i above and right of x_j .



We must have put x_i so high for a reason: There must be some x_k ($k < i$) with $x_k <_P x_i$ and x_i was therefore assigned a y -coordinate barely above x_k . In the drawing above, x_k is therefore on one of the horizontal lines, either to the left or to the right of x_j . It cannot be on the right (i.e. on the blue line and $j < k < i$) since that would imply $x_j < x_k$ (otherwise we would have made a mistake already earlier, but i was minimal) and because of $x_k < x_i$ and transitivity we would also have $x_j < x_i$ contradicting our assumption of $x_j \parallel x_i$.

So we have that x_k is to the left of x_j (i.e. on the red line and $k < j$). This means $x_k \parallel x_j$ and therefore $x_j \cdot \downarrow_{x_k}^{x_i}$ is a copy of $\cdot \downarrow$ in P . Since in L we have $x_k < x_j < x_i$ it is not ordered by L , a contradiction. \square

Higher dimensional posets. There is a natural way to generalize the notion from the previous few pages to arbitrary dimension. First note that for $d \in \mathbb{N}$ the set \mathbb{R}^d becomes an (infinite) poset via the *dominance order* which simply means that for $x = (x_1, \dots, x_d) \in \mathbb{R}^d$, $y = (y_1, \dots, y_d) \in \mathbb{R}^d$ we have

$$x \leq_{\text{dom}} y :\Leftrightarrow \forall i \in [d]: x_i \leq y_i.$$

We say a poset P is d -dimensional if it is a subposet of $(\mathbb{R}^d, \leq_{\text{dom}})$ and not a subposet of $(\mathbb{R}^{d-1}, \leq_{\text{dom}})$.

Theorem 5.8. *A poset P is d -dimensional if and only if d is smallest number of linear extensions of P whose intersection is P .*

Proof. It suffices to show that P is a subposet of \mathbb{R}^d if and only if $P = \bigcap_{i=1}^d L_i$ for some linear extensions L_1, \dots, L_d of P .

“ \Rightarrow ” We can assume, without loss of generality, that no two points share a coordinate (we always break ties without changing the relations in the poset). Then define L_i as the order of the elements on the projection of P to the i -th coordinate. This is a set of linear extensions with intersection P .

“ \Leftarrow ” Given linear extensions L_1, \dots, L_d we can take these to assign coordinates. The coordinate of $x \in P$ will be $(r_1, \dots, r_d) \in \mathbb{R}^d$ where r_i is the rank of x in the i -th linear extension, i.e. $r_i = |\{y \in P \mid y \leq_{L_i} x\}|$. It is easy to see that this gives an embedding of P into \mathbb{R}^d . \square

Next we observe that every poset has a well-defined dimension, in other words, every poset can be written as the intersection of d linear orders for some large enough d . We call such a set of linear extensions a *realizer* of P .

We claim that taking all linear extensions $\mathcal{L} = \{L \mid L \text{ is linear extension of } P\}$ works, i.e. $P = \bigcap_{L \in \mathcal{L}} L$. It is clear that “ \subseteq ” holds. We leave the other direction as an exercise.

Our next result shows that the dimension does not only exist but is actually bounded by the width of the poset. We require some preparation though.

Definition 5.9. Let $P = (X, \leq)$ be a poset.

- For $S \subseteq X$ define the *open downset* $D(S) = \{y \in X \mid \exists x \in S: y < x\}$ and the *closed downset* $D[S] = D(S) \cup S$.
- For single element sets $S = \{s\}$ we will simply write $D(s)$ and $D[s]$ instead of $D(\{s\})$ and $D[\{s\}]$.
- In the same way define *open* and *closed upsets*:

$$U(S) = \{y \in X \mid \exists x \in S: x < y\}, \quad U[S] = U(S) \cup S, \\ U(s) = U(\{s\}), \quad U[s] = U[\{s\}].$$

Consider the left of Figure 33 for an example.

Theorem 5.10. *For any poset P its dimension is at most its width, i.e.*

$$\dim(P) \leq w(P).$$

Proof. Our goal is to find a realizer on $w = w(P)$ dimensions.

By Dilworth’s Theorem we can partition P into $w(P)$ chains C_1, \dots, C_w .

Claim. For any chain C there is a linear extension L of P such that for any $x \in C$ and $x \parallel y$ we have $y <_L x$.

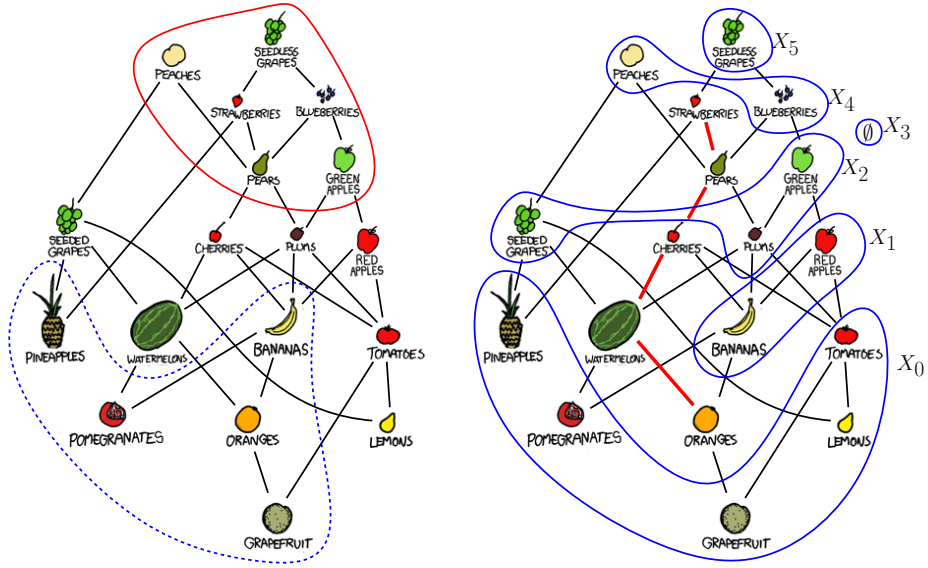


Figure 33: On the left: The open upset of $\{\text{Cherries, Red Apples, Blueberries}\}$ is shown in red and the closed downset of $\{\text{Pineapples Bananas}\}$ is shown in dashed blue.

On the right: Sketch for the claim from Theorem 5.10. Consider the chain $C = \{\text{Oranges, Watermelons, Cherries, Pears, Strawberries}\}$. Then L is a linear extension that puts the elements from C as late as possible, for instance like this (elements from C highlighted): $L : \text{Grapefruit, Lemons, Tomatoes, Pineapples, Pomegranates, Oranges, Bananas, Red Apples, Watermelons, Plums, Green Apples, Seeded Grapes, Cherries, Pears, Peaches, Blueberries, Strawberries, Seedless Grapes}$

Note first that once we have proved this claim, we have proved the theorem since the linear extensions L_1, \dots, L_w we get for the chains C_1, \dots, C_w are a realizer of P , meaning $P = \bigcap_{i=1}^w L_i$. It is clear that the right side is an extension of the left side. Now consider if $x \parallel y$. Then $x \in C_i$ for some $i \in [w]$, and $y \in C_j$ for some $j \neq i$. Then we have $y <_{L_i} x$ and $x <_{L_j} y$ so neither relation occurs in $L_i \cap L_j$.

Think of L_i as a linear extension of P that puts the elements from C_i as late as possible. A sketch is given on the right of Figure 33.

Proof of claim. Denote the elements of C by $x_1 < x_2 < \dots < x_k$ and consider their upsets. Note that clearly $U(x_1) \subset U(x_2) \subset \dots \subset U(x_k)$.

Now define $X_0 := X \setminus U[x_1]$, $X_j := U(x_j) \setminus U[x_{j+1}]$ (for $1 \leq j < k$) and $X_k := U(x_k)$ and let $P_j := (X_j, \leq)$ be the poset induced by X_j ($0 \leq j \leq k$).

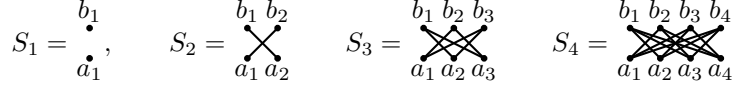
Given a linear extension L_j for each P_j ($0 \leq j \leq k$) we now define the linear extension L of P as:

$$L : L_0 \ x_1 \ L_1 \ x_2 \ L_2 \ \dots \ x_k \ L_k.$$

First, convince yourself that L really is a linear extension of P , i.e. every element occurs exactly once and if $x <_P y$ then x occurs before y .

Now assume $x \in C$ and $x \parallel y$. Say $x = x_l$. Then $y \notin U(x_l)$ so y is not part of any X_j for $j \geq l$. This means $y <_L x_l = x$ as claimed. \square

Example 5.11 (Standard Examples). For a positive integer n define the poset $S_n = (\{a_1, \dots, a_n, b_1, \dots, b_n\}, \leq)$ where $a_i \leq b_j \Leftrightarrow i \neq j$ and no (non-reflexive) relations within $\{a_1, \dots, a_n\}$ and $\{b_1, \dots, b_n\}$ respectively. For instance:



We call these posets the *standard examples*. We claim that $\dim(S_n) = n = w(S_n)$ (for $n \geq 2$), so the standard examples make the bound from the last Theorem tight.

To see that $\dim(S_n) \geq n$, consider a realizer of S_n (consisting of linear extensions). Since $a_i \parallel b_i$ for $i \in [n]$, some linear extension L in the realizer must have $b_i <_L a_i$. Since every remaining b_j is greater than a_i they must all go right of a_i and every remaining a_j is smaller than b_i , so they must all go left of b_i . This means that $b_j <_L a_j$ for no $j \neq i$. To reverse all pairs $(a_i, b_i)_{i \in [n]}$ therefore requires n linear extensions.

5.2 Capturing Posets between two Lines

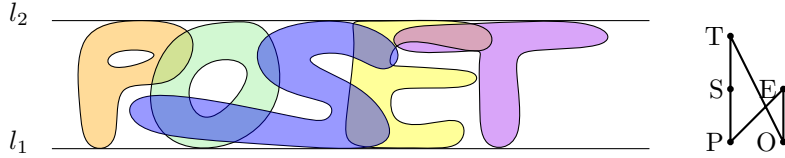
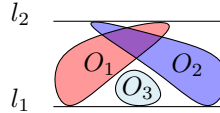


Figure 34: On the left we drew five shapes between two horizontal lines. From this we obtain a poset (shown on the right) by considering a shape less than another shape if it is left of it.

Imagine l_1 and l_2 are two horizontal lines in the plane. An object $O \subseteq \mathbb{R}^2$ is *spanned* between l_1, l_2 if it is contained in the strip in between l_1 and l_2 , is connected and touches both l_1 and l_2 .

With a set of such objects we associate a poset where the strict relation is given as “left of”, meaning that $O_1 \leq O_2$ if and only if $O_1 \cap O_2 = \emptyset$ and the leftmost point of O_1 is left of the leftmost point of O_2 . This clearly gives an order relation (once we add the reflexive relations). Note that since the objects touch the top and bottom line we avoid situations like this



where one might be inclined to say that O_1 is left of O_3 is left of O_2 but O_1 is not left of O_2 . Since this is forbidden our “left of” is really consistent with all reasonable intuitions one might have about leftness.

In the following we ask: What kind of posets can be represented in such a way and what shapes are required?

Observation 5.12 (Straight Lines). If P is 2-dimensional then P can be represented by straight segments spanned between the two lines.

To see this, assume $P = L_1 \cap L_2$ for two linear orders L_1, L_2 . Then place the elements of P on l_1 in the order given by L_1 , also place them onto l_2 in the order given by L_2 and connect the points corresponding to the same element. This gives a line segment s_x for each $x \in X$.

Now observe:

$$x <_P y \Leftrightarrow (x <_{L_1} y \text{ and } x <_{L_2} y) \Leftrightarrow s_x \text{ left of } s_y.$$

The reverse holds as well, i.e. any poset represented by straight segments spanned between l_1 and l_2 is at most two-dimensional and a realizer is given by sorting the segments according to their endpoints on l_1 and l_2 .

The second type of object we consider are axis aligned rectangles. Note that since those rectangles must be tangential to l_1 and l_2 , they are already uniquely determined by their leftmost and rightmost x -coordinate. In fact, the setting is not “really” two-dimensional and is easily seen to be equivalent to the setting of *interval orders* where:

An interval order $P = (X, \leq)$ is given by a set X of open bounded intervals in \mathbb{R} with $(a, b) <_P (c, d)$ iff $b <_{\mathbb{R}} c$.

We remark that not all interval orders are 2-dimensional but we will not prove this until later. We first characterize them in terms of a forbidden subposet:

Theorem 5.13 (Fishburn). $P = (X, \leq)$ is an interval order if and only if $2 \oplus 2 \not\subseteq P$, i.e. there is no copy of the poset \mathbb{I} contained in P .

Proof. “ \Rightarrow ” Assume there is a copy of $2 \oplus 2$ labeled like this.

$$\begin{array}{cc} b & d \\ \uparrow & \uparrow \\ a & c \end{array}$$

If P could be written as an interval order, then a, b, c, d would be represented by intervals (a_l, a_r) , (b_l, b_r) , (c_l, c_r) , (d_l, d_r) with the following properties:

- $a_r <_{\mathbb{R}} b_l$ (since $a <_P b$)
- $b_l <_{\mathbb{R}} c_r$ (since $c \parallel_P b$)
- $c_r <_{\mathbb{R}} d_l$ (since $c <_P d$)
- $d_l <_{\mathbb{R}} a_r$ (since $a \parallel_P d$)

This requires $a_r <_{\mathbb{R}} b_l <_{\mathbb{R}} c_r <_{\mathbb{R}} d_l <_{\mathbb{R}} a_r$, which is clearly not possible.

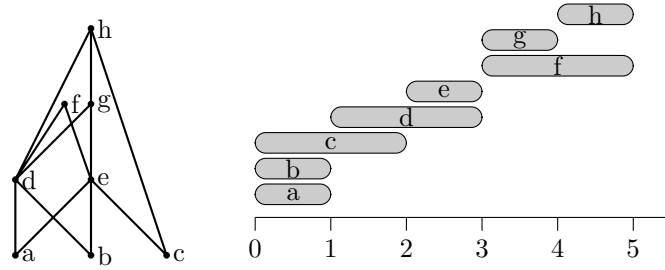
“ \Leftarrow ” Let P be $(2 \oplus 2)$ -free. The crucial observation is that for any two elements $b, d \in X$ we have $D(b) \subseteq D(d)$ or $D(b) \supseteq D(d)$. Assume not: Then $D(b) \not\subseteq D(d)$ and $D(b) \not\supseteq D(d)$. This implies $b \parallel d$ and the existence of $a \in D(b) \setminus D(d)$ and $c \in D(d) \setminus D(b)$. Convince yourself that this means that a, b, c, d forms a copy of \mathbb{I} (check $a < b$, $c < d$, $a \parallel c$, $a \parallel d$, $b \parallel c$) contradicting our assumption.

Given the observation, we can order the downsets $\mathcal{D} = \{D(x) \mid x \in X\}$ by strict inclusion, i.e. $\emptyset = D_0 \subset D_1 \subset D_2 \subset \dots \subset D_\mu$.

For $x \in X$ we choose the interval (a_x, b_x) such that

$$D(x) = D_{a_x}, \quad b_x = \begin{cases} \min\{\beta \mid x \in D_\beta\} & \text{if } x \notin \max(P) \\ \mu + 1 & \text{if } x \in \max(P). \end{cases}$$

Observe that $a_x < b_x$. Also, if $x <_P y$ then $x \in D(y)$ so $b_x \leq a_y$ and conversely, if $x \parallel y$ then $x \notin D(y)$ so $b_x \geq a_y$ and also $y \notin D(x)$ so $b_y \geq a_x$. This means $(a_x, b_x) \cap (a_y, b_y) \neq \emptyset$ and neither interval is left of the other. \square



$$\mathcal{D} = \left\{ \underbrace{D_0 = \emptyset}_{D(a)=D(b)=D(c)}, \underbrace{D_1 = ab}_{D(d)}, \underbrace{D_2 = abc}_{D(e)}, \underbrace{D_3 = abcde}_{D(f)=D(g)}, \underbrace{D_4 = abcdeg}_{D(h)} \right\}$$

Figure 35: Example for the construction in Fishburn's Theorem. The poset on the left contains no $2 \oplus 2$. We determine all downsets $D_0 \subseteq \dots \subseteq D_4$. We pick the interval orders as the Theorem suggests, for instance, the interval for d is $(1, 3)$ since D_1 is the downset of d and D_3 is the first downset containing d .

We now consider the case where the objects are (open) triangles spanned between l_1 and l_2 with the base on l_1 and tip on l_2 (see pictures below).

Theorem 5.14. *P is a triangle order if and only if there is a linear extension of P that orders all copies of $2 \oplus 2$.*

Here we say a $2 \oplus 2$ is ordered by a linear order L , if L puts the element of one chain both before the elements of the other chain. In other words, if $2 \oplus 2$ is labeled like this:

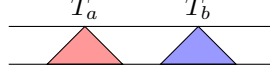
$$\begin{array}{cc} b & d \\ \uparrow & \uparrow \\ a & c \end{array}$$

then we want either $a <_L b <_L c <_L d$ or $c <_L d <_L a <_L b$.

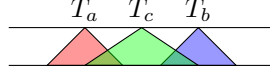
Proof. “ \Rightarrow ” If P is a triangle order then we claim that taking L as the order of the tips of the triangles from left to right works. Note that L is a total order (obviously) and a linear extension of P : If a triangle is left of another triangle ($T_1 <_P T_2$), then in particular its tip is left of the other triangle ($T_1 <_L T_2$).

Now consider a copy of $2 \oplus 2$ with the same labeling as above.

Then a, b, c, d correspond to triangles T_a, T_b, T_c, T_d and T_a is left of T_b .



Assume for contradiction that the tip of T_c is in between the tips of T_a and T_b . Since $c \parallel b$ we need T_c to intersect T_b , maybe like this:



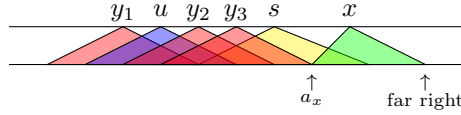
Now there is no way to add T_d : It must be right of T_c but not right of T_a which is clearly impossible.

“ \Leftarrow ” Assume $L : x_1 < x_2 < \dots < x_n$ is a linear extension of P that orders all copies of $2 \oplus 2$. We argue that P is a triangle order. We construct a triangle T_x for every $x \in P$, given by the horizontal position of the three points, namely the tip t_x and base (a_x, b_x) .

The tips should simply be ordered according to L , so $t_{x_i} = i$ for each $i \in [n]$. We pick the bases of the triangles one after the other, going through the elements of P from left to right according to L .

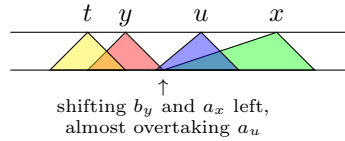
For x_1 we define its base arbitrarily. Having defined triangles for x_1, \dots, x_k , $k \geq 1$, we now show how to add T_x for $x = x_{k+1}$ with the desired relationships to the already existing triangles.

Consider the downset $D(x)$. We choose $a_x = \max\{b_y \mid y \in D(x)\}$, so as far left as possible while still right of every triangle that T_x should be right of. The right end point b_x is chosen somewhere far right.



In the picture we assume $D(x) = \{y_1, y_2, y_3\}$, so $a_x = \max\{b_{y_1}, b_{y_2}, b_{y_3}\} = b_{y_3}$. This ensures that $x >_P y$ implies that T_x is right of T_y (for any y). Also (by choice of the tips) T_x is not left of an existing triangle. However, the pictures shows that there might be $u \notin D(x)$ such that T_x is right of T_u . We need to fix this.

We do this by shifting a_x to the left. But remember that a_x needs to be right of b_y for each $y \in D(x)$, so as soon as a_x would overtake such a b_y , we just take it with us, shifting it to the left as well. Now we need to argue that this does not cause T_y to be suddenly left of some T_u for $y \not\leq u$. So assume a_x and b_y are about to overtake the left endpoint a_u of T_u , maybe like this:



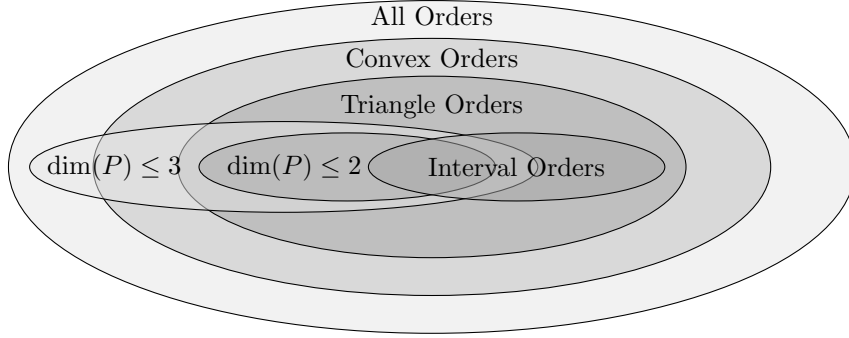
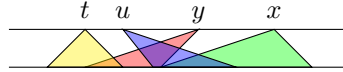


Figure 36: Venn diagram showing the relationships between some types of orders we consider here and in the following.

Now remember: The reason we are shifting is that there is some $t \not\leq x$ with T_t still left of T_x . Since we almost overtook a_u already, this triangle T_t is also left of T_u so $t < u$ (otherwise we would have made a mistake earlier). Combining what we know ($y < x, t < u, y \not\leq u, t \not\leq x$) we see that y, x, t, u forms a copy of $2 \oplus 2$ like this:

$$\begin{array}{cc} x & u \\ \uparrow & \uparrow \\ y & t \end{array}$$

Since x is the most recent element, it comes last in L , and since L orders every copy of $2 \oplus 2$ we have $t <_L u <_L y <_L x$ and the picture above was actually misleading. In truth we are in this situation:



Here, shifting a_x and b_y past a_u is no problem as T_u and T_y will still intersect. So we can shift as far as we need which proves the claim. \square

As an overview over the results of this section consider the Venn diagram in Figure 36. The class of convex and triangle orders are already defined geometrically as orders arising from convex sets or triangles spanned between lines, respectively. But remember that interval orders and orders of dimension at most two also arise from shapes in that way: From axis aligned boxes and segments. In Theorem 5.15 we will furthermore see that the set of all orders arises from y -monotone curves.

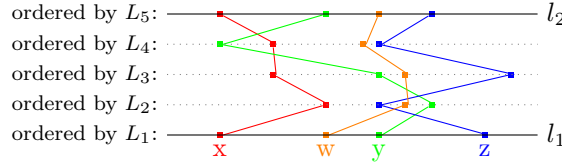
The containments are easy to see: Any straight line is a (degenerated) triangle and every triangle is convex. To see the inclusion of interval orders in triangle orders, observe that if we only allow “boring” triangles that have the tip above their base, then triangles intersect iff their bases intersect, so these triangles behave like intervals. Alternatively, this relation immediately follows from Theorem 5.13 and Theorem 5.14.

In Lemma 5.16 we show that there is indeed a non-convex three dimensional order. We remark without proof that the “small” circle of interval orders already contains posets of arbitrary dimension: The interval order defined by all

closed intervals with end points in $[n]$ has dimension at least $\log \log n + (\frac{1}{2} + o(1)) \log \log \log n$.

Theorem 5.15. *Every poset P can be represented by y -monotone curves spanned between l_1 and l_2 .*

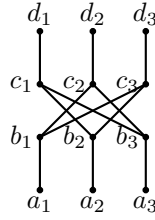
Proof. Take any realizer of P , i.e. $P = L_1 \cap L_2 \cap \dots \cap L_k$ for linear orders L_1, \dots, L_k . Assume without loss of generality $k \geq 2$ and introduce $k-2$ lines in between l_1 and l_2 . This gives k lines in total. On the i -th line we distribute the element of P in increasing order according to L_i . We then connect all points belonging to the same element with a straight segments.



Now it is obvious that $x <_{L_i} y$ for each i if and only if the line for x is left of the line for y . \square

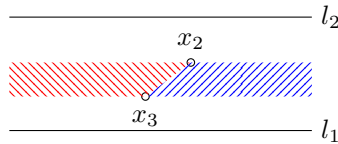
Lemma 5.16. *There is a 3-dimensional order that is not a convex order.*

Proof. Here is one such poset. It is a modified standard example:



Assume it can be represented by convex shapes $C_{a_1}, C_{a_2}, \dots, C_{d_3}$ that are spanned between l_1 and l_2 . Note that this implies that each such shape C_x contains a straight line S_x that is spanned between l_1 and l_2 .

Since $a_i \parallel d_i$ for $i = 1, 2, 3$ the shapes C_{a_i} and C_{d_i} intersect. So take a point $x_i \in C_{a_i} \cap C_{d_i}$. Assume without loss of generality that x_1 has the middle y -coordinate. With respect to the straight line from x_2 to x_3 , we have that x_1 is either left of it (red area) or right of it (blue) area:

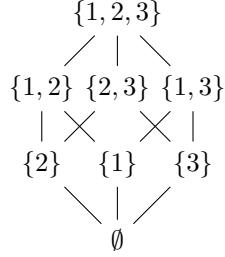


Assume x_1 is on the left. Then since $c_1 > a_2$ and $c_1 > a_3$, the shape C_{c_1} must be right of x_2 and x_3 (since $x_2 \in C_{a_2}$, $x_3 \in C_{a_3}$), but since $c_1 < d_1$ it must also be left of x_1 (since $x_1 \in C_{d_1}$). The same holds for the straight line S_{c_1} . Clearly such a straight line does not exist.

If x_1 is on the right instead, we run into the same problem with the line S_{b_1} : It has to be right of x_1 but left of x_2 and x_3 . \square

5.3 Sets of Sets and Multisets – Lattices

We now consider posets whose elements are sets ordered by inclusion meaning $A \leq B$ iff $A \subseteq B$ (we will prefer the latter notation). Usually the sets we consider are subsets of $[n]$ and the most important example is the *Boolean lattice* B_n , the family of *all* subsets of $[n]$. Take for instance B_3 :



Two sets are incomparable if they are pairwise not included in one another. For instance $\{1\} \parallel \{2, 3\}$ and $\{1, 2\} \parallel \{1, 3\}$. Note also that $\{\{1, 2\}, \{2, 3\}, \{1, 3\}\}$ is a largest antichain in B_3 .

The next Theorem generalizes this observation.

Theorem 5.17 (Sperner's Theorem). $w(B_n) = \binom{n}{\lceil \frac{n}{2} \rceil}$.

Proof. Note that two different sets of the same size are never included in one another. So taking all k -subsets of $[n]$ (i.e. all subsets of size k) yields an antichain of size $\binom{n}{k}$. Choosing $k = \lceil \frac{n}{2} \rceil$ maximizes its size and proves the lower bound.

To prove the upper bound, we introduce a new notion: For a permutation π of $[n]$ we say that π *meets* a set $A \subseteq [n]$ if the elements of A form a prefix of π , meaning:

$$A = \{\pi(1), \pi(2), \dots, \pi(|A|)\}.$$

For instance, the permutation $\pi = 24513$ meets $\{2, 4\}$ and $\{1, 2, 4, 5\}$ but does not meet $\{4, 5\}$.

Now consider a antichain \mathcal{F} . We count the number of permutations that meet an element of \mathcal{F} . Note that the sets met by a single permutation π form a chain, so no π can meet several elements of \mathcal{F} .

So clearly the sets $\{\pi \mid \pi \text{ meets } A\}$ are disjoint for different $A \in \mathcal{F}$ and we have:

$$\sum_{A \in \mathcal{F}} |\{\pi \mid \pi \text{ meets } A\}| \leq n!$$

For any given $A \subseteq [n]$ there are $|A|!(n - |A|)!$ permutations that meet A : We know that the first $|A|$ elements of such a π are given by A and can be arranged in $|A|!$ ways, and the remaining $n - |A|$ elements can be arranged in $(n - |A|)!$

ways. So we have:

$$\begin{aligned}
& \sum_{A \in \mathcal{F}} |A|!(n - |A|)! \leq n! \\
\Leftrightarrow & \sum_{A \in \mathcal{F}} \frac{1}{\binom{n}{|A|}} \leq 1 \\
\Rightarrow & \sum_{A \in \mathcal{F}} \frac{1}{\binom{n}{\lceil \frac{n}{2} \rceil}} \leq 1 \\
\Leftrightarrow & |\mathcal{F}| \leq \binom{n}{\lceil \frac{n}{2} \rceil}.
\end{aligned}$$

In the first step we divided both sides by $n!$, then we used that $k = \lceil \frac{n}{2} \rceil$ maximizes $\binom{n}{k}$ and thus made the term under the sum independent of A . \square

Remark. Note that if \mathcal{F} actually is a largest antichain in B_n , then all inequalities used in the above proof must be tight. In particular, each $\binom{n}{|A|}$ must actually have been equal to $\binom{n}{\lfloor \frac{n}{2} \rfloor} = \binom{n}{\lceil \frac{n}{2} \rceil}$. So only sets of size $\lfloor \frac{n}{2} \rfloor$ or $\lceil \frac{n}{2} \rceil$ are used in \mathcal{F} . From this it is easy to see that:

- If n is even, then $\mathcal{F} = \binom{[n]}{\frac{n}{2}}$ is the unique largest antichain.
- If n is odd, then all largest antichains are contained in $\binom{[n]}{\lfloor \frac{n}{2} \rfloor} \cup \binom{[n]}{\lceil \frac{n}{2} \rceil}$.

Next we consider so-called intersecting families. We say $\mathcal{F} \subseteq 2^n$ is *intersecting* if $\forall A, B \in \mathcal{F}: A \cap B \neq \emptyset$.

Observation 5.18. If $\mathcal{F} \subseteq 2^{[n]}$ is intersecting, then $|\mathcal{F}| \leq 2^{n-1}$ and this bound can be attained.

Proof. For any element $A \in \mathcal{F}$, the complement $[n] \setminus A$ is disjoint from A and cannot be in \mathcal{F} . So taking complements is an injection that maps elements from \mathcal{F} to elements not in \mathcal{F} . Therefore $|\mathcal{F}| \leq 2^n - |\mathcal{F}|$ and thus $|\mathcal{F}| \leq 2^{n-1}$.

To attain this bound, consider $\mathcal{F}_x = \{A \subseteq [n] \mid x \in A\}$ for some fixed $x \in [n]$. Clearly this is an intersecting family of size 2^{n-1} . \square

So this problem has a fairly straightforward and boring solution. But what about the maximum cardinality of an intersecting k -family, i.e. we only allow sets of size k ? If $k = 1$, then we can clearly only pick one set (any two different sets of size 1 do not intersect). If $k = 2$, then it is easy to see that $n - 1$ is best possible using the sets $\{1, 2\}, \{1, 3\}, \{1, 4\}, \dots, \{1, n\}$. Is the best strategy still the “obvious” one in general? It turns out the answer is “yes”, but the proof is a bit more involved.

Theorem 5.19 (Erdős-Ko-Rado). *For two integers n and $0 < k \leq \frac{n}{2}$ the maximum size of an intersecting k -family of $[n]$ is $\binom{n-1}{k-1}$.*

Note that we have the requirement of $2k \leq n$ since the problem becomes trivial otherwise: Two sets of size bigger than $n/2$ always intersect.

Proof. The lower bound is easy, just take $\mathcal{F}_x^k = \{A \subseteq [n] \mid x \in A, |A| = k\}$, for some fixed $x \in [n]$. Clearly, \mathcal{F}_x^k is an intersecting family of the desired size.

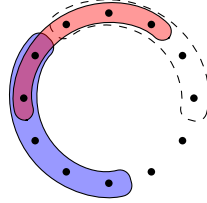
For the upper bound we use a similar approach as in Sperner's Theorem. Recall that a circular permutation σ of $[n]$ is an arrangement of the numbers $1, \dots, n$ on a circle where we do not distinguish between circular shifts, e.g. $53142 \equiv 31425$. The number of circular permutations of $[n]$ is $(n-1)!$.

We say a circular permutation σ *meets* a set $A \subseteq [n]$ if the elements of A appear consecutively on σ . For instance $A = \{1, 2, 4\}$ is met by $\sigma = 25341$ but $B = \{2, 3, 4\}$ is not.

Now fix \mathcal{F} to be an intersecting k -family.

Claim. For any circular permutation σ , the number of sets $A \in \mathcal{F}$ met by σ is at most k .

Proof by Picture. Fix one k -set A_0 met by σ , we draw it in red (here $k = 5$, $n = 12$):



Consider a different k -set $B \in \mathcal{F}$ that is also met by σ (here drawn in blue), it must intersect A_0 . There are $2(k-1)$ choices for such a B , $(k-1)$ for intersecting A_0 at the “clockwise” side and $(k-1)$ for intersecting A_0 at the “counterclockwise” side. But note that if $B \in \mathcal{F}$ intersects some part of A_0 until some gap between two consecutive elements on σ , then the set B' that intersects the opposite part of A_0 starting from that gap cannot be in \mathcal{F} (shown dashed in the picture). This is because $B' \cap B = \emptyset$. Here we use that $2k \leq n$ so B and B' do not intersect on the opposite side either. This shows that there are at most $2(k-1)/2 = k-1$ sets met by σ other than A_0 , so at most k in total. (end claim)

Just like in the proof of Sperner's Theorem, the number of circular permutations meeting $A \in \mathcal{F}$ is $|A|!(n-|A|)! = k!(n-k)!$. We now double count the set

$$\mathcal{S} := \{(A, \sigma) \mid A \in \mathcal{F} \text{ and } \sigma \text{ is circular permutation meeting } A\}.$$

First Way:

$$|\mathcal{S}| = \sum_{\sigma} |\{A \in \mathcal{F} \mid \sigma \text{ meets } A\}| \leq \sum_{\sigma} k = (n-1)! \cdot k.$$

Second Way:

$$|\mathcal{S}| = \sum_{A \in \mathcal{F}} |\{\sigma \mid \sigma \text{ meets } A\}| = \sum_{A \in \mathcal{F}} k!(n-k)! = |\mathcal{F}| \cdot k!(n-k)!$$

Putting this together we obtain:

$$|F| \leq \frac{(n-1)!k}{k!(n-k)!} = \frac{(n-1)!}{(k-1)!(n-k)!} = \binom{n-1}{k-1}. \quad \square$$

5.3.1 Symmetric Chain Partition

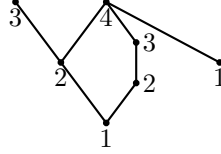
We have already seen in Sperner's Theorem that the width of the Boolean lattice B_n is $\binom{n}{\lceil \frac{n}{2} \rceil}$. Using Dilworth's Theorem, this means B_n can be partitioned into $\binom{n}{\lceil \frac{n}{2} \rceil}$ chains. We are going to prove a version of this that is stronger in two ways. Firstly, we are going to look at a generalization of Boolean lattices. Secondly, we find a particularly interesting partition into chains. The next definition captures what we mean by "interesting".

Definition 5.20.

- Let $P = (X, \leq)$ be a poset. Then the *rank* of $x \in X$ is the size of a largest chain beneath x , i.e.

$$\text{rank}(x) = \max\{|C| \mid C \text{ is a chain with } \max(C) = x\}, \quad x \in X.$$

Consider for instance the following poset in which we annotated each element with its rank:



- Define $\text{rank}(P)$ to be the maximum rank of an element of P (this is just a new name for the height of the poset). The above poset has rank 4.
- A poset P is *ranked* if all maximal chains ending in an element x have the same size. The poset shown above is not ranked since there are maximal chains of size 2, 3 and 4 all ending in the same element. Another way to think about it, is that a poset is ranked iff for any cover relation $x \lessdot y$ the elements x and y have adjacent ranks.
- A chain C is *unrefinable* if there is no $z \in X - C$ with $\min(C) < z < \max(C)$ such that $C \cup \{z\}$ is a chain. Note that if P is ranked, then C is unrefinable if and only if it does not skip any rank.
- A chain C is *symmetric* if it is unrefinable and

$$\text{rank}(\min(C)) + \text{rank}(\max(C)) = \text{rank}(P) + 1.$$

In Boolean lattices this is a very intuitive concept.

- A *symmetric chain decomposition* is a partition of the elements of P into symmetric chains. Figure 37 shows decompositions of some Boolean lattices.

We define $B(m_1, m_2, \dots, m_k)$ to be the poset on all submultisets of $M = \{m_1 \cdot \mathbf{1}, m_2 \cdot \mathbf{2}, \dots, m_k \cdot \mathbf{k}\}$ ordered by inclusion. Here, m_i is the repetition number of the element \mathbf{i} . The elements $\mathbf{1}, \mathbf{2}, \dots, \mathbf{k}$ do not really need to be numbers, we might as well have chosen 🍌 instead of $\mathbf{1}$ and 🍎 instead of $\mathbf{2}$. The elements just need to be distinguishable and numbers happen to be convenient. Recall that if A, B are multisets with types $\mathbf{1}, \dots, \mathbf{k}$ and repetition numbers a_1, \dots, a_k for A and b_1, \dots, b_k for B then by $A \subseteq B$ we mean $a_i \leq b_i$ for all i .

Note that $B(\underbrace{1, \dots, 1}_k) = B_k$ since in that case $M = \{\mathbf{1}, \dots, \mathbf{k}\}$.

Figure 38 depicts $B(1, 2, 1)$, a three dimensional poset that is similar to B_3 except it has an additional “plane” since we can have the element $\mathbf{2}$ not only zero or one times, but also twice.

There are many other ways to explain what these posets are, maybe you prefer one of the following perspectives:

- The poset $B(m_1, \dots, m_k)$ is isomorphic to the poset of all divisors of $p_1^{m_1} \cdot p_2^{m_2} \cdot \dots \cdot p_k^{m_k}$ ordered by divisibility, where p_1, \dots, p_k are distinct primes. This is illustrated for $B(1, 2, 1)$ and $p_1 = 2, p_2 = 3, p_3 = 5$ on the right of Figure 38.
- $B(m_1, \dots, m_k)$ is the product of chains:

$$B(m_1, \dots, m_k) = C_1 \times \dots \times C_k$$

where each C_i is a chain on $m_i + 1$ elements. By the product $P \times Q$ of two posets P and Q we mean the posets with elements $\{(x, y) \mid x \in P, y \in Q\}$ and dominance order, so

$$(x, y) \leq_{P \times Q} (x', y') :\Leftrightarrow x \leq_P x' \text{ and } y \leq_Q y'.$$

- Using particularly easy chains we can view $B(m_1, \dots, m_k)$ as the set $[m_1 + 1] \times [m_2 + 1] \times \dots \times [m_k + 1]$ with dominance order $\leq_{\mathbb{R}^k}$.

Note that $B(m_1, \dots, m_k)$ is ranked. A set $A = \{r_1 \cdot \mathbf{1}, \dots, r_k \cdot \mathbf{k}\}$ has rank $\text{rank}(A) = r_1 + \dots + r_k + 1$ and the rank of the entire poset is $\text{rank}(B(m_1, \dots, m_k)) = m_1 + \dots + m_k + 1$.

Theorem 5.21. $B(m_1, \dots, m_k)$ has a symmetric chain decomposition.

Proof. We do induction on k . If $k = 1$, then $B(m_1)$ is a (symmetric) chain of length $m_1 + 1$. So the trivial partition works.

For $k \geq 2$, consider the subposet $P = B(m_1, \dots, m_{k-1}, 0)$ of $B(m_1, \dots, m_k)$, consisting of those multisets with repetition number 0 for type k . Clearly P is isomorphic to $B(m_1, \dots, m_{k-1})$ and has, by induction hypothesis, a symmetric chain decomposition.

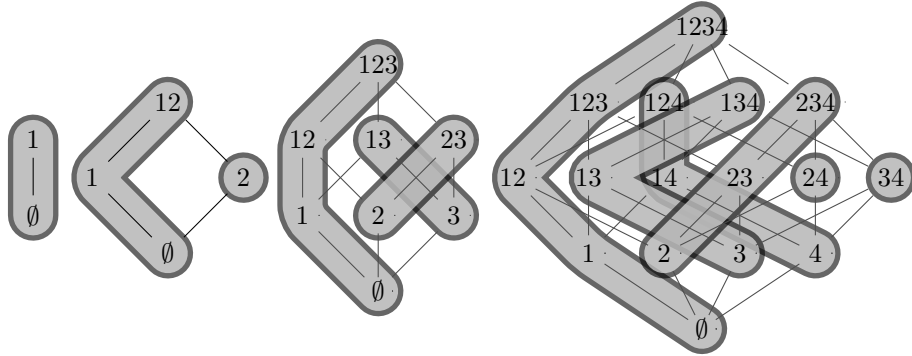


Figure 37: Symmetric chain decompositions of B_1 , B_2 , B_3 and B_4 .

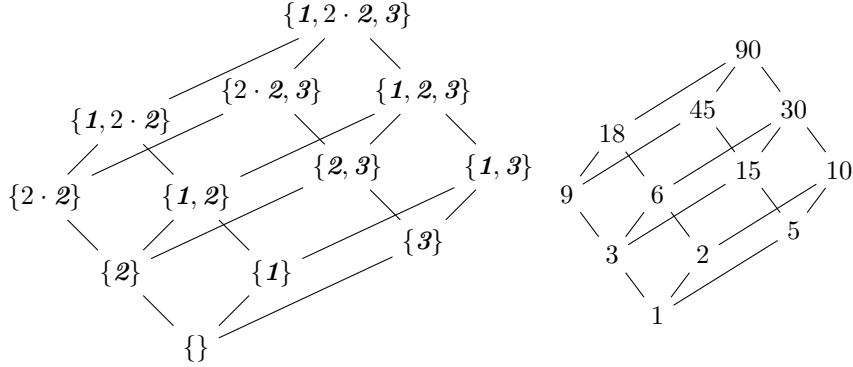


Figure 38: Left: The Hasse Diagram of the poset $B(1, 2, 1)$. Right: Divisors of $90 = 2^1 \cdot 3^2 \cdot 5^1$ ordered by divisibility.

Now the idea is really simple, see Figure 39. Given a symmetric chain composition of P , we first partition $B(m_1, \dots, m_k)$ into “curtains” that run along the chains in P and then we partition the curtains (which are essentially two dimensional grids) into symmetric chains. Now for the formal argument. If $P = C_1 \cup \dots \cup C_R$ is a partition into symmetric chains, then $B(m_1, \dots, m_k) = \text{Cur}_1 \cup \dots \cup \text{Cur}_R$ is a partition into R sets where

$$\text{Cur}_i := \{A \cup \{j \cdot \mathbf{k}\} \mid A \in C_i, 0 \leq j \leq m_k\}.$$

Now focus on one of the chains $C = C_i$ with elements $A_1 \subset A_2 \subset \dots \subset A_l$. The corresponding curtain forms a grid of size $m_k \times l$ standing on its tip as shown in Figure 40. Note that it is not necessarily square. The grid can easily be partitioned into unrefinable chains that look like hooks as shown. If $m_k + 1 = l$ the last hook will be a chain of size 1, otherwise, the last hook will like straight in the picture, either pointing to the top left if $l > m_k + 1$ or to the top right if $l < m_k + 1$. We only need to check that all these “hook-chains” are symmetric.

Consider the first hook

$$A_1 \subset A_2 \subset \dots \subset A_l \subset A_l \cup \{k\} \subset \dots \subset A_l \cup \{m_k \cdot \mathbf{k}\}.$$

Its minimum has rank $|A_1| + 1$ and its maximum has rank $|A_l| + m_k + 1$. Remember that C was symmetric in P so we had $|A_1| + 1 + |A_l| + 1 = \text{rank}(P) + 1$ and get:

$$|A_1| + 1 + |A_l| + m_k + 1 = \text{rank}(P) + 1 + m_k = \text{rank}(B(m_1, \dots, m_k)) + 1.$$

This proves that the first hook is a symmetric chain. Subsequent hooks have their minima at higher ranks but the ranks of the maxima are correspondingly lower so it is easy to see that they are symmetric as well. \square

Figure 37 shows the results of this construction for the Boolean lattices B_1 , B_2 , B_3 and B_4 . You may want to verify your understanding by constructing these symmetric chain partitions yourself. Note that in the case of Boolean lattices all curtains have height 2 and will therefore be partitioned into only one or two hooks.

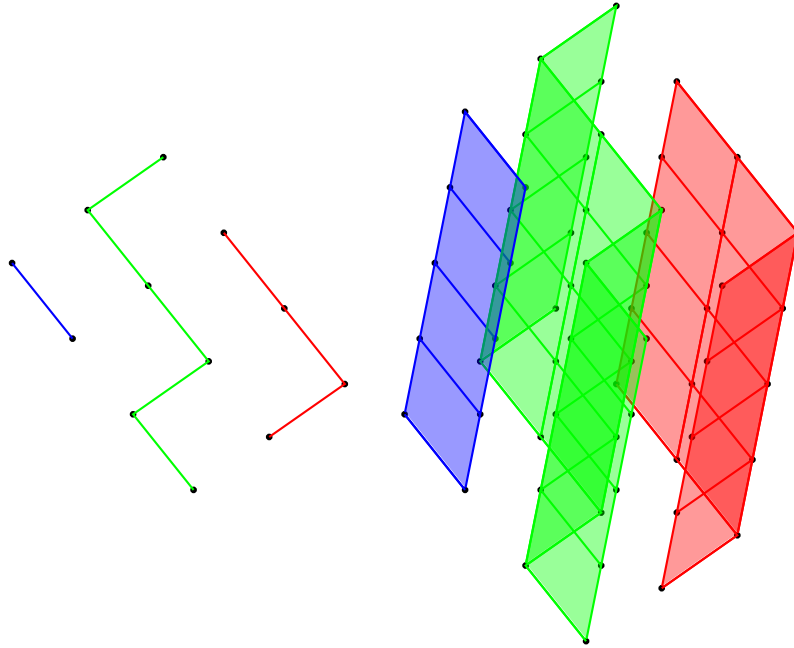
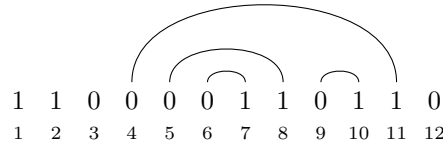


Figure 39: On the left: Partition of $B(2,3)$ into three symmetric chains. On the right: The corresponding partition of $B(2,3,5)$ into three “curtains”.

Since B_n grows exponentially in n , it may be impractical to compute the symmetric chain partition of B_n explicitly for large n . However, there is a simple implicit characterization that allows us to compute for any set $A \subseteq [n]$ what the other sets in the symmetric chain containing A are. We state it here without proof.

The first step is to associate with $A \subseteq [n]$ the characteristic zero-one string c_A that encodes for each $i \in \{1, \dots, 12\}$ in increasing order whether or not i is in A . For $A = \{1, 2, 7, 8, 10, 11\} \subseteq [12]$ this would be $c_A = 110000110110$.

Now, roughly speaking, think of the 0s as opening parenthesis and 1s as closing parenthesis, then c_A is a (not necessarily well-formed) parenthesis expression where some matching pairs can be found. In our example this would be:



The matched positions are those elements that do not vary within the chain. The matched 1s form the minimum m of the chain, here $m = \{7, 8, 10, 11\}$, the matched 0s are the elements missing from the maximum of the chain, so here $M = [12] \setminus \{4, 5, 6, 9\} = \{1, 2, 3, 7, 8, 10, 11, 12\}$. The unmatched positions, here $\{1, 2, 3, 12\}$ will vary within the chain and are added from left to right. So in the symmetric chain partition of B_n the set A will be part of the symmetric

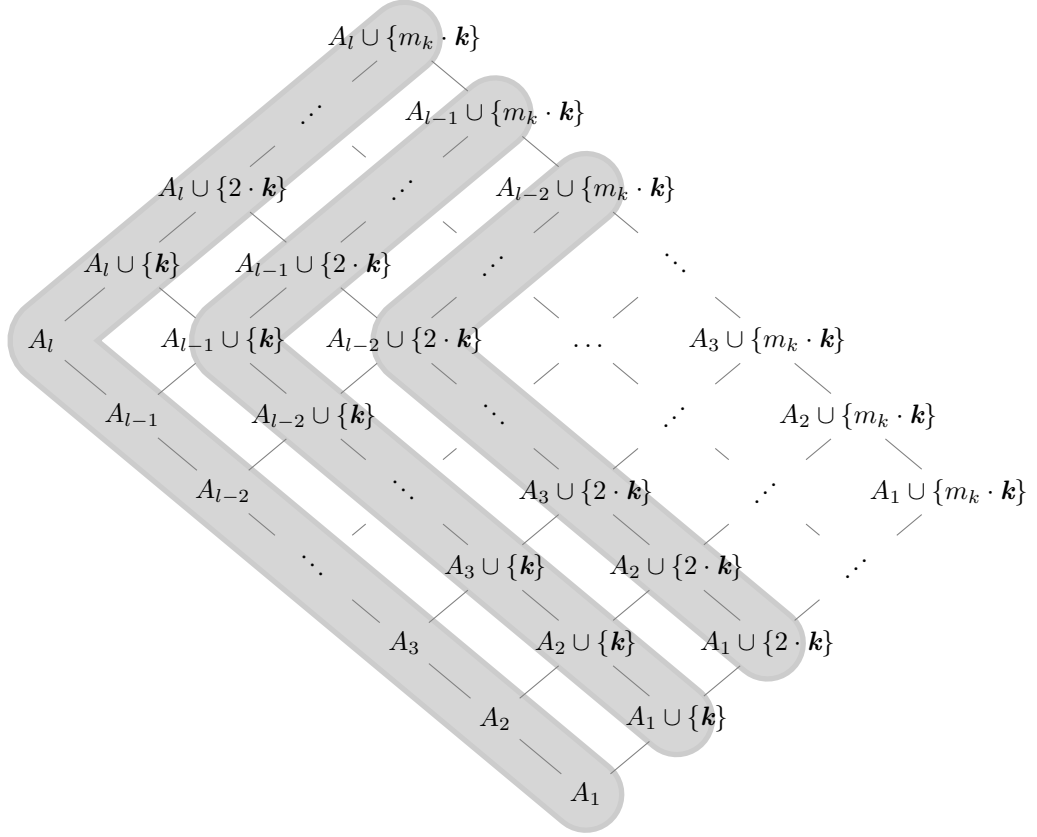


Figure 40: The elements of a curtain can be partitioned into symmetric chains as shown.

chain:

$$\begin{aligned} \{7, 8, 10, 11\} &\subset \{1, 7, 8, 10, 11\} \subset \{1, 2, 7, 8, 10, 11\} \\ &\subset \{1, 2, 3, 7, 8, 10, 11\} \subset \{1, 2, 3, 7, 8, 10, 11, 12\}. \end{aligned}$$

5.4 General Lattices

A poset is a *lattice* if any two elements x, y have a least upper bound, i.e.

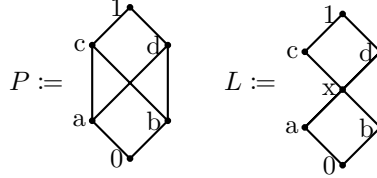
$$\forall x, y \in P: \exists z: (z \geq x \text{ and } z \geq y \text{ and } (\forall z': (z' \geq x \text{ and } z' \geq y) \Rightarrow z \leq z')).$$

We write $z = x \vee y$ and call z the *join* of x and y (it is necessarily unique). We also require that any two elements have a largest lower bound, i.e.

$$\forall x, y \in P: \exists z: (z \leq x \text{ and } z \leq y \text{ and } (\forall z': (z' \leq x \text{ and } z' \leq y) \Rightarrow z \geq z')).$$

We write $z = x \wedge y$ and call z the *meet* of x and y (it is necessarily unique). Note that this implies that lattices have a unique minimum which we call 0 and a unique maximum which we call 1.

The following poset P is not a lattice although it has a maximum and minimum. Note that a and b have no join, the upper bounds of a and b are $\{c, d, 1\}$ but there is no least upper bound (since c and d are incomparable). For similar reasons, c and d have no meet. If we modify P by adding an element x as shown we obtain a lattice L . In it, we have for instance $a \vee b = x$, $c \wedge d = x$.



Note that for two elements x, y of a lattice with $x \leq y$ we always have $x \wedge y = x$ and $x \vee y = y$. This also implies that 1 is the neutral element of the \wedge -operation and 0 is the neutral element of the \vee operation. The Boolean lattices B_n really are lattices where $\wedge = \cap$ and $\vee = \cup$, since $A \cup B$ really is the least set containing A and B and $A \cap B$ is the largest set contained in A and B .

We now consider *Young lattices* $Y(m, n)$. Its elements are Ferrer diagrams with at most m rows and at most n columns ordered by inclusion. Figure 41 shows $Y(2, 3)$.

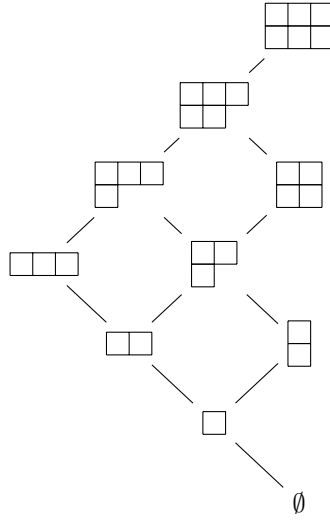


Figure 41: The Young lattice $Y(2, 3)$.

Young lattices are ranked. Diagrams with s squares have rank $s + 1$.

6 Designs

Assume you are the leader of a local brewery and have just invented seven kinds of new beers. You are pretty sure all of them are awesome but are curious whether other people share your opinion. There are some experts, each of which can judge 3 beers (beyond that point they become tipsy and you do not trust their judgment).

You want to make sure that each beer is evaluated in contrast to each other beer, i.e. for each pair of beers there should be one expert that tries both beers. Actually, make sure that there is *exactly* one expert for each pair, since experts are expensive and you don't have any money to waste.

So can you assign beers to the experts and meet this requirement? It turns out your problem has a solution with seven experts as shown in Figure 42.

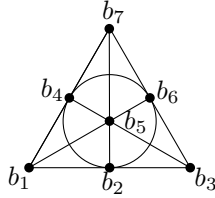


Figure 42: The beers b_1, \dots, b_7 are represented by points. Each set of beers that one expert tries is represented by a straight line or circle. Note that any pair of points uniquely determines a line or circle containing that pair.

In terms of the following definition we have just found a $2-(7, 3, 1)$ -design: $v = 7$ beers, $k = 3$ beers per expert, and each pair ($t = 2$) of beers tasted by $\lambda = 1$ of the experts. This particular design is known as the *Fano plane*.

Definition 6.1. For numbers $t, v, k, \lambda \in \mathbb{N}$, a $t-(v, k, \lambda)$ design with point set V is a multiset \mathcal{B} of sets (“blocks”) of points with

- (i) $|V| = v$,
- (ii) $|B| = k$ for each $B \in \mathcal{B}$,
- (iii) each set of t points is a subset of exactly λ blocks.

A $t-(v, k, \lambda)$ design is also denoted by $S_\lambda(t, k, v)$ and sometimes called *balanced block design*.

Example 6.2.

- Designs with $k = v$, i.e. where each block contains all points, are called *trivial*. This gives a $t-(v, v, |\mathcal{B}|)$ -design for any $t \in [v]$. From now on we assume $k < v$.
- Taking all point sets of size k , i.e. $\mathcal{B} = \binom{[v]}{k}$, yields a $t-(v, k, \binom{v-t}{k-t})$ -design for any $t \in [k]$ (verify this!).
- Consider the point set $V = \mathbb{F}_2^4 \setminus \{\vec{0}\}$, so all bit-strings of length four with addition modulo 2, except for $\vec{0} = (0, 0, 0, 0)$. We have for instance $(1, 1, 0, 0) + (1, 0, 1, 0) = (0, 1, 1, 0)$. Now consider the blocks $\mathcal{B} = \{\{x, y, z\} \mid x, y, z \in \mathbb{F}_2^4, x + y + z = \vec{0}\}$.

Note that any pair of points $x \neq y$ uniquely determines a third point z such that their sum is zero, namely $z = x + y$ (note that $x = -x$ in \mathbb{F}_2^4 so $x \neq -y$ and $z = x + y \neq 0$). So \mathcal{B} is a 2 -($v = 15, k = 3, \lambda = 1$)-design.

6.1 (Non-)Existence of Designs

As a first step we want to understand under which conditions on λ, t, k and v designs with those parameters may exist. In other words we are looking for necessary conditions on the parameters λ, t, k and v .

Theorem 6.3. *If \mathcal{B} is a t -(v, k, λ)-design then*

- (i) *the number of blocks in \mathcal{B} is $|\mathcal{B}| = \lambda \cdot \binom{v}{t} / \binom{k}{t}$,*
- (ii) *if $I \subseteq V$ is a set of size $i \leq t$, then the number of blocks containing I is $r_i = \lambda \cdot \binom{v-i}{t-i} / \binom{k-i}{t-i}$.*

Note that for $i = t$ we get $r_t = \lambda$ as it should be. Also note that the fact that every point is contained in r_1 blocks means that \mathcal{B} is an r_1 -regular hypergraph (you can safely ignore this remark if don't know what hypergraphs are).

Proof. (i) We double count the set

$$\mathcal{S} := \{(T, B) \mid B \in \mathcal{B}, T \subseteq B, |T| = t\}.$$

Firstly, each set $T \subseteq V$ of size t (and there are $\binom{v}{t}$ of those) is contained in exactly λ blocks since that is what a design requires. So $|\mathcal{S}| = \binom{v}{t} \lambda$.

Secondly, each blocks contains $\binom{k}{t}$ subsets of size t , so $|\mathcal{S}| = |\mathcal{B}| \cdot \binom{k}{t}$.

Together we get: $|\mathcal{B}| \cdot \binom{k}{t} = \binom{v}{t} \lambda$ from which the claim follows.

- (ii) Fix an i -set I and double count the set

$$\mathcal{S} := \{(T, B) \mid B \in \mathcal{B}, I \subseteq T \subseteq B, |T| = t\}.$$

Firstly, there are $\binom{v-i}{t-i}$ sets T of size t containing I and for each such T there are λ blocks B containing T . This gives $|\mathcal{S}| = \binom{v-i}{t-i} \cdot \lambda$.

Secondly, for each of the r_I blocks B with $I \subseteq B$ there are $\binom{k-i}{t-i}$ sets T of size t with $I \subseteq T \subseteq B$. This gives $|\mathcal{S}| = r_I \cdot \binom{k-i}{t-i}$.

Together we obtain $\binom{v-i}{t-i} \cdot \lambda = r_I \cdot \binom{k-i}{t-i}$. From this we see that r_I actually only depends on $|I| = i$ and the claim follows. \square

Remark. We consider $t = 2$ to be the default case. Sometimes a 2 -(v, k, λ) design is just called a (v, k, λ) -design. In the other notation the parameter $\lambda = 1$ can be omitted so an $S_1(t, k, v)$ -design is simply a $S(t, k, v)$ -design. In that case we also call it *Steiner System* (hence the “S”).

Corollary 6.4. *If \mathcal{B} is a (v, k, λ) -design (so $t = 2$) the results of the last Theorem become*

- (i) $|\mathcal{B}| = \lambda \frac{v(v-1)}{k(k-1)} = r \frac{v}{k}$ where r is:

$$(ii) \ r = r_1 = \lambda \frac{v-1}{k-1}.$$

Remark 6.5. For any t -(v, k, λ)-design all the numbers we derived above, such as $|\mathcal{B}|, r_i$ ($1 \leq i \leq t$) are integers. In particular, if some choice of parameters t, k, v, λ does not yield integers, no design with those parameters exists.

For example, in any $(v, k = 3, \lambda = 1)$ -design, we have $r_1 = \frac{v-1}{2}$, which is integer only if v is odd, and $|\mathcal{B}| = \frac{(v-1)v}{6}$ which is integer only if $v \in \{0, 1, 3, 4\} \pmod{6}$, so together we need $v \in \{1, 3\} \pmod{6}$. Note that for $v = 3$ we get the trivial design and for $v = 7$ we get the Fano plane, so things seem to fit so far.

However, the necessary conditions for the existence of designs are not yet sufficient. In the following Theorem we derive another necessary condition, this time a lower bound on v .

Theorem 6.6. *In any non-trivial t -($v, k, \lambda = 1$)-design we have $v \geq (t+1)(k-t+1)$.*

Proof. First observe that for $A, B \in \mathcal{B}$ with $A \neq B$ we have $|A \cap B| \leq t-1$. Otherwise some set $T \subseteq |A \cap B|$ of size t would be contained in both A and B , contradicting $\lambda = 1$.

Now choose some $B_0 \in \mathcal{B}$ and a set S of size $t+1$ such that $|S \cap B_0| = t$. This implies that S is not contained in any $B \in \mathcal{B}$ (otherwise the t -set $S \cap B_0$ would be contained in both B_0 and B). Still each of the t -sets $T \subseteq S$ must be contained in some block B_T . Consider two such blocks B_T and $B_{T'}$. Each contains k elements, t of which are in S and $k-t$ are outside of S . The intersection $B_T \cap B_{T'}$ certainly contains $T \cap T'$ which is already of size $t-1$. This means (by the observation in the beginning of the proof) that B_T and $B_{T'}$ must not have any more further intersections outside of S .

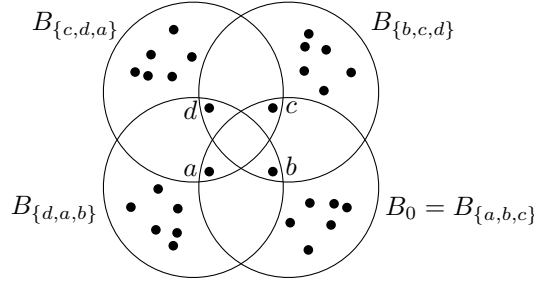


Figure 43: Consider the case $t = 3$ and $k = 9$. After choosing some $B_0 \in \mathcal{B}$ we find a set $S = \{a, b, c, d\}$ of size $t+1$ that intersects B_0 in exactly $t = 3$ points a, b, c . Now no block may contain S (as argued above) but each set $\{a, b, c\}, \{b, c, d\}, \{c, d, a\}, \{d, a, b\}$ must be contained in some block. These four blocks must be disjoint outside of S (they already intersect in $t-1$ elements within S). So we count $k-t = 6$ elements for each of them plus the elements from S , so $v \geq 24 + 4 = 28$.

This means for each of the blocks B_T (and there are $t+1$ of them) that there are $k-t$ elements outside of B_T that are not contained in any other $B_{T'}$.

Together with the elements from S this gives

$$v \geq \underbrace{(t+1)}_{\in S} + \underbrace{(t+1)(k-t)}_{k-t \text{ for each } B_T} = (t+1)(k-t+1). \quad \square$$

So we can conclude, for instance, that no design with parameters $\lambda = 1$, $t = 10$, $v = 72$ and $k = 16$ exists (even though the divisibility conditions of Theorem 6.3 are satisfied).

Remark. In a recent, so far unpublished paper¹⁰ Peter Keevash shows that for any parameters λ, t, k there exist t -(v, k, λ) designs for sufficiently large v . The proof is very sophisticated so we will not include it here.

In the next Theorem, we find a lower bound to the number of blocks of a design.

Theorem 6.7 (Fisher's inequality). *Let \mathcal{B} be a non-trivial (v, k, λ) -design (so $k < v$). Then $|\mathcal{B}| \geq v$.*

Proof. Consider the incidence matrix A with v rows, $|\mathcal{B}|$ columns and values $(a_{p,B})_{p \in V, B \in \mathcal{B}}$ defined as

$$a_{p,B} := \begin{cases} 1, & p \in B, \\ 0, & \text{otherwise.} \end{cases}$$

If a_p, a_q are rows in A , then $a_p \cdot a_q^T$ is the number of blocks containing both p and q . This value is λ for $p \neq q$ and $r = r_1$ for $p = q$. So we have

$$A \cdot A^T = \begin{pmatrix} r & \lambda & \cdots & \cdots & \lambda \\ \lambda & r & & & \lambda \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & r \\ \lambda & \cdots & \cdots & \lambda & r \end{pmatrix}.$$

The determinant is $\det(A \cdot A^T) = rk(r - \lambda)^{v-1}$ (verify this!).

By Corollary 6.4(i) we have $r = \lambda \frac{v-1}{k-1} > \lambda$. So the determinant is not zero and thus the matrix has full rank, i.e. rank v . This means that the matrix A must also have rank at least v . In particular A has at least v columns and therefore $|\mathcal{B}| \geq v$ as claimed. \square

6.2 Construction of Designs

So far we talked about necessary conditions for the existence of designs without actually constructing any – which is our goal in the following. We have to start with a few definitions, though.

Definition 6.8.

- A (v, k, λ) -design with $|\mathcal{B}| = v$ is called *symmetric*. For those designs the matrix A we discussed above is square.

¹⁰<http://arxiv.org/abs/1401.3665> submitted on 15 Jan 2014.

- A symmetric $(v, k, 1)$ -design is called *projective plane*. The Fano plane from Figure 42 is such a projective plane with 7 points and 7 blocks (“lines”).
- A $(v, k = 3, \lambda = 1)$ -design or, equivalently, a Steiner System with parameters $S(t = 2, k = 3, v)$ is called a *Steiner Triple System*.

In the following we write \mathbb{Z}_n for the cyclic group with n elements, in other words, the numbers modulo n .

Definition 6.9. Let $2 \leq k < v$ and $\lambda \geq 1$. A (v, k, λ) -difference set is a set $D = \{d_1, \dots, d_k\} \subseteq \mathbb{Z}_v$ such that each element in $\mathbb{Z}_v \setminus \{0\}$ has multiplicity λ in the multiset $\{d_i - d_j \mid i, j \in [k], i \neq j\}$. In other words, each non-zero number can be written as the difference of two numbers from D in exactly λ ways.

Example 6.10. The set $\{1, 3, 4, 5, 9\} \subseteq \mathbb{Z}_{11}$ is a $(11, 5, 2)$ -difference set, since each of the numbers can be written as a difference in two ways, as follows:

Number	First Way	Second Way
1	$4 - 3$	$5 - 4$
2	$3 - 1$	$5 - 3$
3	$1 - 9$	$4 - 1$
4	$5 - 1$	$9 - 5$
5	$9 - 4$	$3 - 9$
6	$9 - 3$	$4 - 9$
7	$1 - 5$	$5 - 9$
8	$1 - 4$	$9 - 1$
9	$1 - 3$	$3 - 5$
10	$3 - 4$	$4 - 5$

Theorem 6.11. If $D = \{d_1, \dots, d_k\}$ is a (v, k, λ) -difference set then

$$\mathcal{B} := \{D, 1 + D, 2 + D, \dots, v - 1 + D\}$$

is a symmetric (v, k, λ) -design. Here, by $a + D$ we mean the set $\{a + d \mid d \in D\}$.

Proof. Note that there are $k \cdot (k - 1)$ ways to form terms $i - j$ with $i \neq j$ and $i, j \in D$. On the other hand, each non-zero number can be written in exactly λ ways like this, so $\lambda(v - 1) = k \cdot (k - 1)$ which implies $\lambda = \frac{k(k-1)}{v-1} < k$.

From this we conclude that the blocks we proposed above are all distinct: Assume $a + D = b + D$ for $a \neq b$. Then we have a permutation $\pi \in S_k$ such that $a + d_i = b + d_{\pi(i)}$ for all $i \in [k]$. Then $a - b = d_{\pi(i)} - d_i$ for each $i \in [k]$ so we have $k > \lambda$ ways to write $a - b$ as a difference, contradicting the property of the difference set.

By construction we have v points in total and blocks of size k . The only thing left to check is that each pair of two distinct elements $x, y \in \mathbb{Z}_v$ is contained in exactly λ of the blocks. Let $d = x - y$ be the difference of those two elements. Then there are exactly λ ways to pick two elements $x_i, y_i \in D$ with difference d and for each such pair there is a unique shift a such that $x = x_i + a$, $y = y_i + a$. This implies that $x, y \in a + D$ for exactly λ choices of a . \square

Example. If v is a prime power with $v \equiv 3 \pmod{4}$ then

$$\{a^2 \mid a \in \mathbb{Z}_v, a^2 \neq 0\}$$

is a (v, k, λ) -difference set with $k = \frac{v-1}{2}$ and $\lambda = \frac{v-3}{4}$. We do not prove this.

6.3 Projective Planes

We should mention that projective planes actually come from geometry. There, a projective plane of order q is defined as a set of $q^2 + q + 1$ points with a family of subsets of points (lines) such that (i) each line is of size $q + 1$ and (ii) each pair of points is contained in a unique line.

In our setting, projective planes of order q are symmetric $(v, k, \lambda = 1)$ -designs where $k = q + 1$. In that case we can indeed conclude

$$v \stackrel{\text{sym.}}{=} |\mathcal{B}| \stackrel{\text{Cor.6.4}}{=} r \frac{v}{k} \Rightarrow r = k = q + 1$$

so each point is contained in $q + 1$ lines and

$$r \stackrel{\text{Cor.6.4}}{=} \frac{v-1}{k-1} \Rightarrow v = (q+1)q + 1 = q^2 + q + 1,$$

so the number of points matches as well. Before we continue, note the following property:

Claim. Any two lines (blocks) of a projective plane intersect in a unique point.

Proof of Claim. Consider two distinct lines L_1, L_2 and some $x \in L_1 \setminus L_2$. Then consider the family of lines $(L_y)_{y \in L_2}$ where L_y is the unique line containing x and y . These lines are all distinct (otherwise we would find some y_1, y_2 contained in two lines: L_2 and L_{y_1}), so we found $q + 1$ lines that all contain x . This means we must have counted L_1 in the process, otherwise there would be $q + 2$ lines containing x (contradicting $r = q + 1$). We conclude $L_1 \cap L_2 \neq \emptyset$. Clearly $|L_1 \cap L_2| \leq 1$, otherwise we would have a pair of points contained in L_1 and L_2 (contradicting $\lambda = 1$). \square

Construction of Projective Planes

In the following construction of a projective plane of order q we use \mathbb{F}_q to denote a field of size q where q is a prime power.

The points of the plane are equivalence classes of the set $X = \{(x_0, x_1, x_2) \mid x_i \in \mathbb{F}_q, (x_0, x_1, x_2) \neq 0\}$ where two vectors are equivalent if they are scalar multiples of one another. Formally, a point $[x_0, x_1, x_2]$ is the set:

$$[x_0, x_1, x_2] = \{c \cdot (x_0, x_1, x_2) \mid c \in \mathbb{F}_q \setminus \{0\}\}.$$

Note the same point can be described in different ways, for instance $[x_0, x_1, x_2] = [2x_0, 2x_1, 2x_2]$ (for $q \neq 2$). Since $|X| = q^3 - 1$ and $q - 1$ elements represent the same point (are in the same class), we have $\frac{q^3-1}{q-1} = q^2 + q + 1$ points in total, as desired.

For $(a_0, a_1, a_2) \in \mathbb{F}_q^3 \setminus \{\vec{0}\}$ we define the line $L([a_0, a_1, a_2])$ as

$$L([a_0, a_1, a_2]) = \{[x_0, x_1, x_2] \mid a_0x_0 + a_1x_1 + a_2x_2 = 0\}.$$

Note that lines are well defined, that is the definition respects the equivalence classes (either all elements representing a point satisfy the equation of the line or neither of them).

The number of solutions to $a_0x_0 + a_1x_1 + a_2x_2$ is q^2 since, without loss of generality, $a_2 \neq 0$ and thus x_0 and x_1 can be chosen arbitrarily and x_2 is then uniquely determined. Disregarding the solution $x_0 = x_1 = x_2 = 0$ (which does not represent any point) we have $q^2 - 1$ solutions in X that are part of lines, meaning $\frac{q^2-1}{q-1} = q + 1$ points are part of the line, as desired.

Now consider two points $[x_0, x_1, x_2] \neq [y_0, y_1, y_2]$. We show that they are contained in a unique line. Indeed, if $L[a_0, a_1, a_2]$ contains both, then we have $a_0x_0 + a_1x_1 + a_2x_2 = 0$, $a_0y_0 + a_1y_1 + a_2y_2 = 0$. This is a homogeneous system with two equations and three variables a_0, a_1, a_2 , so there exists a solution $(a_0, a_1, a_2) \neq 0$ and all solutions are of the form $c \cdot (a_0, a_1, a_2)$. All those solution triples define the same line, so $L[a_0, a_1, a_2]$ is uniquely determined.

This concludes the construction (and the verification thereof).

Remark. It is conjectured that the order of every projective plane is a prime power, but no proof is known.

Theorem 6.12 (Bruck-Ryser-Chowla-Theorem). *If $v, k, \lambda \in \mathbb{N}$ with $\lambda(v-1) = k(k-1)$ and a symmetric (v, k, λ) -design exists then*

- if v is even, then $k - \lambda$ is a square,
- if v is odd, then $z^2 = (k - \lambda)x^2 + (-1)^{\frac{v-1}{2}}\lambda y^2$ has a solution $(x, y, z) \neq 0$.

6.4 Steiner Triple Systems

We leave projective planes for now and come back to Steiner Triple Systems, i.e. $(v, k = 3, \lambda = 1)$ -designs. From Remark 6.5 we know that $v \in \{1, 3\} \pmod{6}$. The following Theorem shows that this necessary condition on the existence is also sufficient.

Theorem 6.13. *For each $v \in \mathbb{N}$ with $v \in \{1, 3\} \pmod{6}$ a Steiner Triple System with v points exists.*

Proof. The construction is different for $v \equiv 1$ and $v \equiv 3 \pmod{6}$. We start with the latter case.

The points are $V = \mathbb{Z}_{2n+1} \times \mathbb{Z}_3$ so $v = (2n+1) \cdot 3 = 6n+3 \equiv 3 \pmod{6}$. There are two types of blocks:

- type 1: $\{(x, 0), (x, 1), (x, 2)\}$ for each $x \in \mathbb{Z}_{2n+1}$,
- type 2: $\{(x, i), (y, i), (\frac{x+y}{2}, i+1)\}$ for all $x \neq y$ and each $i \in \mathbb{Z}_3$.

We need to show that each pair of elements is contained in a unique block. So consider a pair $(x, i) \neq (y, j)$.

Case 1: $x = y$: Then the pair is contained in the block $\{(x, 0), (x, 1), (x, 2)\}$ (of type 1) and in no other block.

Case 2: $x \neq y, i = j$: The pair is contained in the block $\{(x, i), (y, i), (\frac{x+y}{2}, i+1)\}$ (of type 2) and in no other block.

Case 3: $x \neq y, i \neq j$: Assume without loss of generality that $j = i + 1$ (since $i, j \in \mathbb{Z}_3$ we always have this or $j = i - 1$). Then the pair is contained in the block $\{(x, i), (y', i), (\frac{x+y'}{2}, i + 1)\}$ where $y' = 2y - x$. The pair is contained in no other block.

This concludes the case of $v \equiv 3 \pmod{6}$, we proceed with $v \equiv 1 \pmod{6}$. The construction and verification is more complicated.

As point set, choose $V = (\mathbb{Z}_{2n} \times \mathbb{Z}_3) \cup \{\infty\}$. To simplify notation we write x_i to denote the pairs $(x, i) \in \mathbb{Z}_{2n} \times \mathbb{Z}_3$. The element ∞ is special and we assume $x + \infty = \infty$ for any $x \in V$. Before we define the blocks of the Triple System, we define four types of *base blocks* first:

- $\{0_0, 0_1, 0_2\}$,
- $\{\infty, 0_0, n_1\}, \{\infty, 0_1, n_2\}, \{\infty, 0_2, n_0\}$,
- $\{0_0, x_1, (-x)_1\}, \{0_1, x_2, (-x)_2\}, \{0_2, x_0, (-x)_0\}$, for each $x \in [n - 1]$,
- $\{n_0, x_1, (1 - x)_1\}, \{n_1, x_2, (1 - x)_2\}, \{n_2, x_0, (1 - x)_0\}$, for each $x \in [n]$.

The blocks of the Steiner Triple System are

$$\mathcal{B} = \{a_0 + B \mid a \in \{0, 1, \dots, n - 1\}, B \text{ is base block}\}.$$

where the addition of elements is defined coordinate wise, in particular $a_0 + x_i = (a + x)_i$.

We need to check that there is a unique block containing u, v with $u \neq v$. We do not consider all cases here, just two important ones.

Case 1: $u = \infty, v = (x, i)$: If $x \leq n - 1$, then the block is $x_0 + \{\infty, 0_i, n_{i+1}\}$. If $x \geq n$ then we find $(x - n)_0 + \{\infty, 0_{i-1}, n_i\}$.

Case 2: $u = x_i, v = y_i$: Without loss of generality $x < y$. Either $y - x = 2s$ (the difference is even), then x_i, y_i is contained in the block $(y - s)_0 + \{0_{i-1}, s_i, (-s)_i\}$. Or $y - x = 2s - 1$ then x_i, y_i is contained in the block $(y - s)_0 + \{n_{i-1}, s_i, (1 - s)_i\}$. $\dots \square$

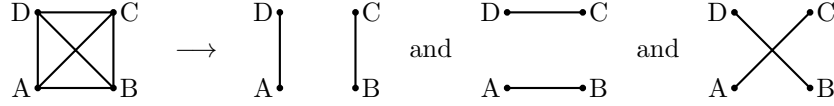
6.5 Resolvable Designs

Consider a game played by three players (like skat). Assume there is a tournament for this game, with n players. Each set of three players should play exactly λ times against each other. There are several time slots during which the games are scheduled. Of course every player can only play one game at a time. In an optimal schedule all players have a game in each time slot. In the sense of the following definition, such a schedule corresponds to a resolvable $(n, 3, \lambda)$ -design, where parallel classes correspond to the time slots.

Definition 6.14. A *parallel class* of a (v, k, λ) -design is a set of disjoint blocks forming a partition of the point set V .

A partition of \mathcal{B} into parallel classes is called *resolution*. A design is *resolvable* if it has a resolution.

Note that projective planes are examples for non-resolvable designs since no disjoint blocks exist. An example for a resolvable $(v = 4, k = 2, \lambda = 1)$ -design with the corresponding parallel classes is shown in the next picture, where the points are A, B, C, D and the blocks are depicted by edges.



This is a special case of a $(v = q^2, k = q, \lambda = 1)$ -design. Such designs are called affine planes and are, as we show now, always resolvable.

Theorem 6.15. *Any $(v = q^2, k = q, \lambda = 1)$ -design is resolvable.*

Proof. The main observation is the following:

Claim. For each block B and $x \notin B$ there is a unique block B' with $x \in B'$ and $B \cap B' = \emptyset$.

Proof of Claim. For each $y \in B$ there is a unique block B_y such that $\{x, y\} \in B_y$. These blocks are distinct (since $|B_y \cap B| \leq 1$) and all contain x . Because of $r \stackrel{\text{Cor. 6.4}}{=} \lambda \frac{v-1}{k-1} = \frac{q^2-1}{q-1} = q+1$ there is exactly one block left that contains x that is different from each B_y . It is disjoint from B . \square

With the claim proved, everything falls into place. Start with a maximal set of pairwise disjoint blocks. This set forms a parallel class (this is not obvious, but easy to prove). Then, in the next phase, take some other set of disjoint blocks not yet considered. This forms a parallel class in the same way. Continue like this until all blocks are handled (it is easy to verify that this works). \square

6.6 Latin Squares

In the following we consider a structure that seems unrelated to designs at first, but the final Theorem will establish an interesting connection to affine planes.

A *Latin square* of order $n \geq 1$ is an $n \times n$ array filled with numbers from the cyclic group $\mathbb{Z}_n = \{0, \dots, n-1\}$ such that each row and each column contains each number from \mathbb{Z}_n exactly once. Consider for instance the following two Latin squares

$$A = \begin{array}{c|c|c|c} 1 & 0 & 2 & 3 \\ \hline 0 & 1 & 3 & 2 \\ \hline 2 & 3 & 1 & 0 \\ \hline 3 & 2 & 0 & 1 \end{array} \quad B = \begin{array}{c|c|c|c} 1 & 0 & 3 & 2 \\ \hline 3 & 2 & 1 & 0 \\ \hline 0 & 1 & 2 & 3 \\ \hline 2 & 3 & 0 & 1 \end{array}.$$

We denote the set of positions containing the number i in a Latin square L by $L(i)$, in the example above we have for instance $B(3) = \{(1, 3), (2, 1), (3, 4), (4, 2)\}$. Another way to think about Latin squares is to say that they are a partition $L(0) \cup \dots \cup L(n-1)$ of $[n] \times [n]$ such that each set $L(i)$ are a set of positions such that rooks (from the game chess) placed onto those positions cannot attack each other. Every solution to a Sudoku puzzle is a Latin square with $n = 9$, although the reverse is not true: Latin squares do not respect the constraint for the 3×3 boxes.

There are many ways to transform a Latin square to similar Latin squares: You could permute the rows or columns or “rename” the numbers (swap $A(1)$

and $A(0)$ for instance). At least the latter operation seems uninteresting as renaming numbers yields the same partition, just with different labels. To capture the idea of “entirely different Latin squares, we propose the notion of *orthogonality* defined in the following.

Let A, B be Latin squares of order n with entries $A = (a_{ij})_{i,j}$, $B = (b_{ij})_{i,j}$. The *juxtaposition* of A and B is the $n \times n$ array where each position simply contains the corresponding numbers of A and B , i.e.

$$A \otimes B = ((a_{ij}, b_{ij}))_{i,j} \text{ with entries in } \mathbb{Z}_n \times \mathbb{Z}_n.$$

For instance, with A and B from the example above we get

$$A \otimes B = \begin{array}{c|c|c|c} (1,1) & (0,0) & (2,3) & (3,2) \\ \hline (0,3) & (1,2) & (3,1) & (2,0) \\ \hline (2,0) & (3,1) & (1,2) & (0,3) \\ \hline (3,2) & (2,3) & (0,0) & (1,1) \end{array}.$$

We say A, B are *orthogonal* if $A \otimes B$ contains each pair from $\mathbb{Z}_n \times \mathbb{Z}_n$ exactly once. The two Latin squares above are not orthogonal because some pairs occur twice (e.g. $(2, 0)$) and, necessarily, some pairs therefore occur not at all (e.g. $(1, 3)$).

An example for two Latin squares A and C with A orthogonal to C is

$$A = \begin{array}{c|c|c|c} 1 & 0 & 2 & 3 \\ \hline 0 & 1 & 3 & 2 \\ \hline 2 & 3 & 1 & 0 \\ \hline 3 & 2 & 0 & 1 \end{array}, \quad C = \begin{array}{c|c|c|c} 1 & 0 & 2 & 3 \\ \hline 3 & 2 & 0 & 1 \\ \hline 0 & 1 & 3 & 2 \\ \hline 2 & 3 & 1 & 0 \end{array}, \quad A \otimes C = \begin{array}{c|c|c|c} (1,1) & (0,0) & (2,2) & (3,3) \\ \hline (0,3) & (1,2) & (3,0) & (2,1) \\ \hline (2,0) & (3,1) & (1,3) & (0,2) \\ \hline (3,2) & (2,3) & (0,1) & (1,0) \end{array}.$$

Note that this notion of orthogonality has little to do with geometric notions orthogonality you may be familiar with. Observe the following:

- If A is orthogonal to B , then B is orthogonal to A .
- If A is orthogonal to B , then “renaming” numbers in A or B (for instance replace every 0 with a 1 and vice versa) preserves orthogonality.

Remark. For $n \in \{2, 6\}$ there are no two orthogonal Latin squares of order n . For $n = 2$ this is easy to see since the only Latin squares with $n = 2$ are

$$L_1 = \begin{array}{c|c} 0 & 1 \\ \hline 1 & 0 \end{array} \quad \text{and} \quad L_2 = \begin{array}{c|c} 1 & 0 \\ \hline 0 & 1 \end{array}.$$

They are not orthogonal to each other nor to themselves (only for $n = 1$ can a Latin square be orthogonal to itself). For $n = 6$ the argument is non-trivial.

For $n \in \mathbb{N} \setminus \{2, 6\}$, a pair of orthogonal Latin squares of order n exists. The proof of this is highly non-trivial, in fact it took over a hundred years to show that there is a pair of orthogonal Latin squares of order 10.

Our goal in the following is to construct a large set A_1, \dots, A_k of Latin squares of order n that are *MOLS* (*mutually orthogonal Latin squares*), meaning A_i is orthogonal to A_j for $i \neq j$.

Theorem 6.16. *Let n be a positive integer, $r \in [n-1]$ non-zero and co-prime to n , i.e. $\gcd(n, r) = 1$. Then $L_n^r = (r \cdot i + j \pmod{n})_{i,j}$ is a Latin square.*

To clarify how L_n^r looks, consider the example $n = 5$ and $r = 2$. “Going right” corresponds to $+1$ and going down corresponds to $+2$ which gives

$$L_5^2 = \begin{array}{c|c|c|c|c} 3 & 4 & 0 & 1 & 2 \\ \hline 0 & 1 & 2 & 3 & 4 \\ \hline 2 & 3 & 4 & 0 & 1 \\ \hline 4 & 0 & 1 & 2 & 3 \\ \hline 1 & 2 & 3 & 4 & 0 \end{array}.$$

Proof. We have to show that each row and column contains each number exactly once. For rows this is clear, since the i -th row contains $r \cdot i + 1, r \cdot i + 2, \dots, r \cdot i + n$, which traverses all numbers modulo n . The column j contains $r + j, 2r + j, \dots, n \cdot r + j$, all modulo n . Assume two of those numbers are identical, say $i_1 \cdot r + j \equiv i_2 \cdot r + j \pmod{n}$ we conclude $(i_1 - i_2) \cdot r \equiv 0 \pmod{n}$. Since $\gcd(n, r) = 1$ (meaning r has an inverse modulo n) we actually had $i_1 = i_2$. So no column contains a number twice and L_n^r really is a Latin square of order n . \square

Theorem 6.17. *If n is prime, then L_n^1, \dots, L_n^{n-1} are $n - 1$ MOLS of order n .*

Proof. Since n is prime, $\gcd(n, i) = 1$ for all $i \in [n - 1]$ so by Theorem 6.16 L_n^1, \dots, L_n^{n-1} are Latin squares of order n .

Now consider two of those squares L_n^r and L_n^s with $r \neq s$. We need to show that they are orthogonal, so suppose some pair of numbers from $\mathbb{Z}_n \times \mathbb{Z}_n$ appears in $L_n^r \otimes L_n^s$ in positions (i, j) and (k, l) . By definition of L_n^r and L_n^s this gives the identity:

$$(r \cdot i + j, s \cdot i + j) = (r \cdot k + l, s \cdot k + l).$$

So $r \cdot (i - k) = l - j = s \cdot (i - k)$ which implies $(r - s)(i - k) = 0$. The numbers modulo n form a field and a product can only be zero if one of the factors is zero. Since $r \neq s$ we obtain $i = k$ and therefore also $l = j$. In particular we showed that the same pair cannot appear in *distinct* positions and of $L_n^r \otimes L_n^s$, so L_n^r is orthogonal to L_n^s as desired. \square

Remark. For a prime power $n = p^k$ there is a field $\mathbb{F}_n = \{\alpha_0 = 0, \alpha_1, \dots, \alpha_{n-1}\}$ of order n . For it we can define $L_n^{\alpha_r} = (\alpha_r \cdot \alpha_i + \alpha_j)_{i,j}$ and generalize Theorem 6.16 and Theorem 6.17 accordingly.

Lemma 6.18. *For any $n \geq 2$ there are at most $n - 1$ MOLS of order n .*

Proof. Assume we have a set A_1, \dots, A_k of MOLS. As orthogonality is not affected by “renaming” the numbers in a Latin square, we may assume without loss of generality that each A_i has $0 \mid 1 \mid 2 \mid \dots \mid n - 1$ as its first row. Now consider the entry at position $(2, 1)$. If A is orthogonal to B then $a_{21} \neq b_{21}$ because all pairs $(0, 0), \dots, (n - 1, n - 1)$ already appear in row 1 of $A \otimes B$. So the entries in $(2, 1)$ must be mutually distinct. They must also be non-zero since the first column of each A_i already contains zero in position $(1, 1)$. Therefore, we started with at most $k \leq n - 1$ Latin squares. \square

And now for the theorem that establishes the connection to designs. We will only prove the equivalence of (i) and (iv), but we will do so constructively.

Theorem 6.19. *For $n \geq 2$ each of the following is equivalent:*

- (i) *There exist $n - 1$ mutually orthogonal Latin squares of order n .*

(ii) There exists a finite field of order n .

(iii) n is a prime power, i.e. $n = p^k$.

(iv) There exists a $(v = n^2, k = n, \lambda = 1)$ -design (an affine plane).

Proof. (i) \Rightarrow (iv). Let A_1, \dots, A_{n-1} be $n - 1$ MOLS of order n . We need to construct a $(v = n^2, k = n, \lambda = 1)$ -design. Recall (from Corollary 6.4) that the number of blocks in such a design is necessarily $|\mathcal{B}| = n(n + 1)$ and from Theorem 6.15 that the design is necessarily resolvable, so it splits into $n + 1$ parallel classes of n disjoint blocks each (this resolution will be apparent).

The points of the design are $[n] \times [n]$, so the set of positions in Latin squares of order n .

Recall how each Latin square A is a partition of the positions $[n] \times [n] = A(0) \cup \dots \cup A(n-1)$ where $A(i)$ are the positions containing number $i \in \mathbb{Z}_n$. These will be blocks of the design. In addition, each row $R(i) = \{(i, j) \mid j \in [n]\}$ and each column $C(j) = \{(i, j) \mid i \in [n]\}$ is a block. So in total the blocks are:

$$\mathcal{B} = \{A_r(s) \mid r \in [n-1], s \in \mathbb{Z}_n\} \cup \{R(i) \mid i \in [n]\} \cup \{C(j) \mid j \in [n]\}.$$

We need to show that each pair of distinct points (i, j) and (k, l) appears in exactly one block. We first show that they are contained in *at most* one block.

Case 1: $i = k$. Both points are contained in the i -th row so both are in $R(i)$. They cannot be contained in the same column (otherwise they would not be distinct) and they are not both contained in any $A_r(s)$ since that would mean that the Latin square A_r contains the number s twice in the i -th row. In particular, no block other than $R(i)$ contains both points.

Case 2: $j = l$. Similar to Case 1, the points are contained in $C(j)$ and in no other block.

Case 3: $i \neq k, j \neq l$. The points are not in the same row or column so in no $R(i)$ or $C(j)$. But assume (for contradiction) that we have $(i, j), (k, l) \in A_r(s) \cap A_t(u)$ for $(r, s) \neq (t, u)$. Since the blocks originating from the same Latin square are disjoint (since those blocks form a partition), we have $r \neq t$. So in A_r there is the number s in both positions (i.e. at (i, j) and (k, l)) and in A_t there is u in both positions. This means $A_r \otimes A_t$ has (s, u) in both positions, contradicting the fact that A_r and A_t are orthogonal. So each pair of positions is contained in at most one block.

To see that each pair of positions is contained in *at least* one block we double count the set:

$$\mathcal{S} = \{((i, j), (k, l), B) \mid B \in \mathcal{B}, (i, j) \neq (k, l), B \text{ contains } (i, j) \text{ and } (k, l)\}.$$

$$\begin{aligned} \text{Firstly, } |\mathcal{S}| &= \sum_{(i,j) \neq (k,l)} \{B \mid B \text{ contains } (i,j) \text{ and } (k,l)\} \\ &\leq \sum_{(i,j) \neq (k,l)} 1 = \binom{n^2}{2}. \quad (\text{uses "at most one"}) \end{aligned}$$

$$\begin{aligned} \text{Secondly, } |\mathcal{S}| &= \sum_B \{((i,j), (k,l)) \mid (i,j) \neq (k,l), B \text{ contains } (i,j) \text{ and } (k,l)\} \\ &= \sum_B \binom{|B|}{2} = (n^2 + n) \binom{n}{2} = \binom{n^2}{2}. \end{aligned}$$

So we realize that the “ \leq ” is actually an equality so in particular each pair of points is contained in exactly one block.

(iv) \Rightarrow (i) Assume we have a $(v = n^2, k = n, \lambda = 1)$ -design. As argued before (Corollary 6.4), there are $|\mathcal{B}| = (n+1) \cdot n$ blocks in total. Since affine planes are resolvable by Theorem 6.15 we get $n+1$ parallel classes $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_{n+1}$ consisting of n blocks each. We use two of these parallel classes \mathcal{B}_1 and \mathcal{B}_2 to find “coordinates” for the n^2 points and the remaining $n-1$ parallel classes to define MOIS.

We label the blocks from \mathcal{B}_1 and \mathcal{B}_2 as $\mathcal{B}_1 = \{R(1), R(2), \dots, R(n)\}$ and $\mathcal{B}_2 = \{C(1), C(2), \dots, C(n)\}$. For $i, j \in [n]$, we know that every pair of blocks $R(i)$ and $C(j)$ intersects in exactly one point p_{ij} . Because of $t = 2$ and $\lambda = 1$, they cannot intersect in two points and because of the claim from the proof of Theorem 6.15 they cannot be disjoint. This means that the blocks from \mathcal{B}_1 and \mathcal{B}_2 intersect as shown in Figure 44 so

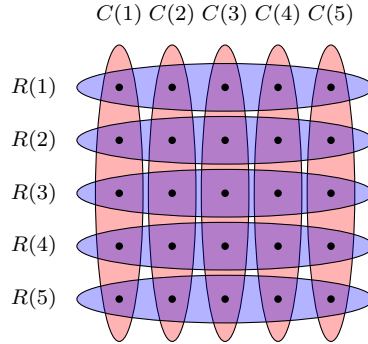


Figure 44: The two different parallel classes \mathcal{B}_1 and \mathcal{B}_2 intersect like this when arranging the points accordingly.

we can interpret the blocks $R(i)$ as rows and $C(j)$ as columns ($i, j \in [n]$). To simplify notation, identify the point p_{ij} of the design with the position $(i, j) \in [n] \times [n]$ in Latin squares. Then for each $l \in \{3, \dots, n+1\}$ we interpret the n blocks $B_l(0), B_l(2), \dots, B_l(n-1)$ of the parallel class \mathcal{B}_l as a Latin square where the positions containing a number $r \in \mathbb{Z}_n$ are given by $B_l(r)$. This really is a valid Latin square since each $B_l(r)$ intersects each $R(i)$ and $C(j)$ in exactly one point so each number occurs in exactly one row and column.

It is also easy to verify that the $n - 1$ Latin squares we get from \mathcal{B}_3 to \mathcal{B}_{n+1} are mutually orthogonal, since for any distinct $s, t \in \{3, \dots, n + 1\}$ and $r_1, r_2 \in \mathbb{Z}_n$ we know that $|B_s(r_1) \cap B_t(r_2)| = 1$ which just means that the pair $(r_1, r_2) \in \mathbb{Z}_n \times \mathbb{Z}_n$ occurs exactly once in the juxtaposition of the Latin square for \mathcal{B}_s and \mathcal{B}_t . \square

References

- [Kas61] Pieter W. Kasteleyn. The statistics of dimers on a lattice: I. the number of dimer arrangements on a quadratic lattice. *Physica*, 27(12):1209–1225, 1961.
- [TF61] H. N. V. Temperley and Michael E. Fisher. Dimer problem in statistical mechanics—an exact result. *Philos. Mag. (8)*, 6:1061–1063, 1961.